

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:11:27 ; Search time 260.2 seconds
(without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-21

Perfect score: 55
Sequence: 1 VLNAVSTSPLE 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	96	US-60-160-209-4252	Sequence 4252, Ap
2	55	100.0	102	US-60-160-203-4414	Sequence 4414, Ap
3	55	100.0	261	US-09-245-764-7	Sequence 7, Appl
4	55	100.0	374	US-09-702-021-14	Sequence 14, Appl
5	55	100.0	399	US-09-05882-1128	Sequence 1128, Ap
6	36	65.5	104	US-09-107-532-4325	Sequence 4325, Ap
7	36	65.5	292	US-09-429-755-315	Sequence 315, App
8	36	65.5	292	US-09-443-686-532	Sequence 532, App
9	36	65.5	292	US-09-483-672A-532	Sequence 532, App
10	36	65.5	292	US-09-534-825A-315	Sequence 315, App

11	36	65.5	292	US-09-536-857-532	Sequence 532, App
12	36	65.5	292	US-09-593-793A-532	Sequence 532, App
13	36	65.5	292	US-09-636-215-532	Sequence 532, App
14	36	65.5	292	US-09-651-236-532	Sequence 532, App
15	36	65.5	292	US-09-657-279-532	Sequence 532, App
16	36	65.5	292	US-09-679-426-532	Sequence 532, App
17	36	65.5	292	US-09-685-166-532	Sequence 532, App
18	36	65.5	292	US-09-699-295-315	Sequence 315, App
19	36	65.5	292	US-09-709-729-532	Sequence 532, App
20	36	65.5	292	US-09-759-143-532	Sequence 532, App
21	36	65.5	292	US-09-780-669-532	Sequence 532, App
22	36	65.5	292	US-09-810-936-315	Sequence 315, App
23	36	65.5	329	US-09-288-946-376	Sequence 376, App
24	36	65.5	329	US-09-289-198-299	Sequence 299, App
25	36	65.5	329	US-09-352-616A-376	Sequence 376, App
26	36	65.5	329	US-09-429-755-299	Sequence 299, App
27	36	65.5	329	US-09-443-686-376	Sequence 376, App
28	36	65.5	329	US-09-483-672A-376	Sequence 376, App
29	36	65.5	329	US-09-534-825A-299	Sequence 299, App
30	36	65.5	329	US-09-536-857-376	Sequence 376, App
31	36	65.5	329	US-09-593-793A-376	Sequence 376, App
32	36	65.5	329	US-09-636-215-376	Sequence 376, App
33	36	65.5	329	US-09-651-236-376	Sequence 376, App
34	36	65.5	329	US-09-657-279-376	Sequence 376, App
35	36	65.5	329	US-09-679-426-376	Sequence 376, App
36	36	65.5	329	US-09-679-426-376	Sequence 376, App
37	36	65.5	329	US-09-685-166-376	Sequence 376, App
38	36	65.5	329	US-09-699-295-299	Sequence 299, App
39	36	65.5	329	US-09-709-729-376	Sequence 376, App
40	36	65.5	329	US-09-759-143-376	Sequence 376, App
41	36	65.5	329	US-09-780-669-376	Sequence 376, App
42	36	65.5	329	US-09-810-936-299	Sequence 299, App
43	36	65.5	329	US-09-417-507-24307	Sequence 24307, A
44	36	65.5	329	US-09-107-433-3049	Sequence 3049, App
45	36	65.5	188	US-09-248-796-27815	Sequence 27815, A

ALIGNMENTS

RESULT 1

US-60-160-209-4252

Sequence 4252, Application US/60160209

GENERAL INFORMATION:

APPLICANT: BONAZZI, VIVIAN

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES

TITLE OF INVENTION: THEROP

FILE REFERENCE: C1000113

CURRENT APPLICATION NUMBER: US/60/160,209

CURRENT FILING DATE: 1999-10-19

NUMBER OF SEQ ID NOS: 4646

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4252

LENGTH: 96

TYPE: PRT

ORGANISM: HUMAN

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(96)

OTHER INFORMATION: Xaa - Any Amino Acid

US-60-160-209-4252

Query Match 100.0%; Score 55; DB 23; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNAVSTSPLE 12
DB 7 VLNAVSTSPLE 18

RESULT 2
US-60-160-203-4414
; Sequence 4414, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VYIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4414
; LENGTH: 102
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-4414

Query Match 100.0%; Score 55; DB 23; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNASTVSPLLE 12
DB 15 VLNASTVSPLLE 26

RESULT 3
US-09-245-764-7
; Sequence 7, Application US/09245764
; GENERAL INFORMATION:
; APPLICANT: Hogarth, P. Mark
; APPLICANT: Powell, Maree S.
; APPLICANT: McKenzie, Ian F.C.
; APPLICANT: Maxwell, Kelly F.
; APPLICANT: Garrett, Thomas P.J.
; APPLICANT: Epa. Vidana
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES AND MODELS OF FC RECEPTORS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4102-4
; CURRENT APPLICATION NUMBER: US/09/245,764
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/099,994
; EARLIER FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: 60/073,972
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-245-764-7

Query Match 100.0%; Score 55; DB 16; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNASTVSPLLE 12
DB 174 VLNASTVSPLLE 185

RESULT 4
US-09-702-021-14
; Sequence 14, Application US/09702021
; GENERAL INFORMATION:
; APPLICANT: C. Fraser
; TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF
; FILE REFERENCE: 7853-217

; CURRENT APPLICATION NUMBER: US/09/702,021
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-021-14

Query Match 100.0%; Score 55; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNASTVSPLLE 12
DB 193 VLNASTVSPLLE 204

RESULT 5
PCT-US00-05882-1128
; Sequence 1128, Application PC/TUS0005882
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: PA106PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05882
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1128
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (208)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (349)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1128

Query Match 100.0%; Score 55; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNASTVSPLLE 12
DB 218 VLNASTVSPLLE 229

RESULT 6
US-09-107-532-4325
; Sequence 4325, Application US/09107532
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
; NUMBER OF SEQUENCES: 7308
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA

ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER:
OPERATING SYSTEM:
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085598
FILING DATE: May 14, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4325:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...104
US-09-107-532-4325

Query Match 65.5%; Score 36; DB 15; Length 104;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 NASVSPLE 12
|||:|:|:|
Db 72 NASITEPFE 81

RESULT 7
US-09-429-755-315
Sequence 315, Application US/09429755A
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Smith, John M.
APPLICANT: Reed, Steven G.
APPLICANT: Misher, Lynda
APPLICANT: Retter, Marc W.
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.419C6
CURRENT APPLICATION NUMBER: US/09/429,755A
CURRENT FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 315
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 315
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-429-755-315

Query Match 65.5%; Score 36; DB 18; Length 292;
Best Local Similarity 70.0%; Pred. No. 1,1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 NASVSPLE 12
|||:|:|:|
Db 259 SASVSPLE 268

RESULT 8
US-09-443-686-532
Sequence 532, Application US/09443686
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9A
CURRENT APPLICATION NUMBER: US/09/443,686
CURRENT FILING DATE: 1999-11-18
NUMBER OF SEQ ID NOS: 551
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-443-686-532

Query Match 65.5%; Score 36; DB 18; Length 292;
Best Local Similarity 70.0%; Pred. No. 1,1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 NASVSPLE 12
|||:|:|:|
Db 259 SASVSPLE 268

RESULT 9
US-09-483-672A-532
Sequence 532, Application US/09483672A
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Solk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Wang, Aijun
APPLICANT: Meagher, Madeline
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.42711C11
CURRENT APPLICATION NUMBER: US/09/483,672A
CURRENT FILING DATE: 2000-01-14
NUMBER OF SEQ ID NOS: 590
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens

US-09-483-672A-532

Query Match

Best Local Similarity 65.5%; Score 36; DB 18; Length 292;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;OY 3 NASVSPLE 12
DB 259 SASIVSPLE 268

RESULT 10

US-09-534-825A-315
Sequence 315, Application US/09534825A
GENERAL INFORMATION:
APPLICANT: Fridakis, Tony N.
APPLICANT: Smith, John M.
APPLICANT: Reed, Steven G.
APPLICANT: Misher, Lynda
APPLICANT: Retter, Marc W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
FILE REFERENCE: 210121.419C7
CURRENT APPLICATION NUMBER: US/09/534, 825A
CURRENT FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 317
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 315
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-534-825A-315

Query Match

Best Local Similarity 65.5%; Score 36; DB 19; Length 292;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;OY 3 NASVSPLE 12
DB 259 SASIVSPLE 268

RESULT 11

US-09-536-857-532
Sequence 532, Application US/09536857
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.4271C12
CURRENT APPLICATION NUMBER: US/09/536, 857
CURRENT FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 592
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-536-857-532

Query Match

Best Local Similarity 65.5%; Score 36; DB 19; Length 292;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;OY 3 NASVSPLE 12
DB 259 SASIVSPLE 268

RESULT 12

US-09-593-793A-532
Sequence 532, Application US/09593793A
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.4271C15
CURRENT APPLICATION NUMBER: US/09/593, 793A
CURRENT FILING DATE: 2000-06-13
NUMBER OF SEQ ID NOS: 814
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-593-793A-532

Query Match

Best Local Similarity 65.5%; Score 36; DB 19; Length 292;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;OY 3 NASVSPLE 12
DB 259 SASIVSPLE 268

RESULT 13

US-09-636-215-532
Sequence 532, Application US/09636215
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.42717C17
CURRENT APPLICATION NUMBER: US/09/636,215
CURRENT FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 852
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-636-215-532

Query Match 65.5%; Score 36; DB 20; Length 292;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 NASVTSPLLE 12
:|:|:|:|:|
Db 259 SASIVSPLLE 268

RESULT 14
US-09-651-236-532
Sequence 532, Application US/09651236
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.42718C18
CURRENT APPLICATION NUMBER: US/09/651,236
CURRENT FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 865
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-651-236-532

Query Match 65.5%; Score 36; DB 20; Length 292;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 NASVTSPLLE 12
:|:|:|:|:|
Db 259 SASIVSPLLE 268

RESULT 15
US-09-657-279-532
Sequence 532, Application US/09657279
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C19
CURRENT APPLICATION NUMBER: US/09/657,279
CURRENT FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 877
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-657-279-532

Query Match 65.5%; Score 36; DB 20; Length 292;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 NASVTSPLLE 12
:|:|:|:|:|
Db 259 SASIVSPLLE 268

Search completed: June 4, 2001, 12:20:01
Job time: 514 sec

Tue Jun 5 07:08:17 2001

us-09-284-107-21.rapm

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:11:38 ; Search time 84.28 seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-21
Perfect score: 55
Sequence: 1 VLNAVSTSPLE 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA.New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	12	5	US-09-284-107-21
2	36	65.5	292	5	US-09-590-583-315
3	36	65.5	292	5	US-09-822-827-532
4	36	65.5	292	5	US-09-825-301-4
5	36	65.5	292	5	US-09-439-313-532
6	36	65.5	329	5	US-09-590-583-299
7	36	65.5	329	5	US-09-822-827-376
8	36	65.5	329	5	US-09-825-301-2
9	36	65.5	329	5	US-09-439-313-376
10	34	61.8	356	5	US-09-739-449-8794
11	32	58.2	104	5	US-09-739-449-10537
12	32	58.2	164	5	US-09-270-8498-189198
13	32	58.2	341	5	US-09-270-8498-189198
14	32	58.2	661	5	US-09-232-773A-12
15	32	58.2	687	5	US-09-232-773A-6
16	32	58.2	735	5	US-09-232-773A-14
17	32	58.2	780	5	US-09-232-773A-14
18	32	58.2	1338	5	US-09-426-371-3
19	31	56.4	142	5	US-09-270-8498-181441
20	31	56.4	882	5	US-09-829-671-2
21	30	54.5	30	5	US-09-270-8498-187805
22	30	54.5	44	5	US-09-801-436-344
23	30	54.5	71	5	US-09-270-8498-182930
24	30	54.5	72	5	US-09-270-8498-183902
25	30	54.5	99	5	US-09-270-8498-181440
26	30	54.5	102	1	PCT-US01-01350-117
27	30	54.5	144	5	US-09-640-211A-2260

28	30	54.5	147	5	US-09-270-8498-192928	Sequence 192928,
29	30	54.5	162	5	US-09-270-8498-183119	Sequence 183119,
30	30	54.5	205	1	PCT-US01-11988-1199	Sequence 1199, Ap
31	30	54.5	205	5	US-09-833-245-1199	Sequence 1199, Ap
32	30	54.5	209	5	US-09-270-8498-181024	Sequence 181024,
33	30	54.5	214	5	US-09-739-449-10533	Sequence 10533, A
34	30	54.5	225	5	US-09-640-211A-2277	Sequence 2277, Ap
35	30	54.5	241	1	PCT-US01-04098A-1672	Sequence 1672, Ap
36	30	54.5	249	1	PCT-US01-04098A-3640	Sequence 3640, Ap
37	30	54.5	335	5	US-09-270-8498-182179	Sequence 182179,
38	30	54.5	460	1	PCT-US01-04098A-1317	Sequence 1317, Ap
39	30	54.5	470	1	PCT-US01-04098A-3285	Sequence 3285, Ap
40	30	54.5	551	1	PCT-US01-11988-1202	Sequence 1202, Ap
41	30	54.5	551	5	US-09-833-245-1202	Sequence 1202, Ap
42	30	54.5	665	5	US-09-816-494-2	Sequence 2, Appl
43	30	54.5	672	1	PCT-US00-35017A-1259	Sequence 1259, Ap
44	30	54.5	768	1	PCT-US01-04098A-1011	Sequence 1011, Ap
45	30	54.5	804	1	PCT-US01-04098A-2979	Sequence 2979, Ap

ALIGNMENTS

RESULT 1
US-09-284-107-21
Sequence 21, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
APPLICANT: de Krulff, Cornelis Adriaan John
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: AND PROTEINS HAVING SPECIFIC AFFINITY FOR A TARGET
FILE REFERENCE: 31363200600
CURRENT APPLICATION NUMBER: US/09/284,107
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
PRIOR FILING DATE: 1996-10-08
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 21
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
US-09-284-107-21
US-09-590-583-315
Sequence 315, Application US/09590583
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Reed, Steven G.
APPLICANT: Smith, John M.
APPLICANT: Misher, Linda E.
APPLICANT: Dillon, Davin C.
APPLICANT: Retter, Marc W.
APPLICANT: Wang, Aljun
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
THERAPY AND DIAGNOSIS OF BREAST CANCER

Query Match 100.0%; Score 55; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 VLNAVSTSPLE 12

FILE REFERENCE: 210121.419C9
CURRENT APPLICATION NUMBER: US/09/590.583
CURRENT FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 324
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 315
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-590-583-315

Query Match 65.5%; Score 36; DB 5; Length 292;
Best Local Similarity 70.0%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPILE 12
DB 259 SASIVSPLE 268

RESULT 3
US-09-822-827-532
Sequence 532, Application US/09822827
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.534C1
CURRENT APPLICATION NUMBER: US/09/822.827
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-822-827-532

Query Match 65.5%; Score 36; DB 5; Length 292;
Best Local Similarity 70.0%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPILE 12
DB 259 SASIVSPLE 268

RESULT 4
US-09-825-301-4
Sequence 4, Application US/09825301
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Dillon, David C.
APPLICANT: Molesh, David A.
APPLICANT: Xu, Jiangchun
APPLICANT: Zehentner, Barbara
APPLICANT: Persing, David H.
TITLE OF INVENTION: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION
FILE REFERENCE: 210121.513
CURRENT APPLICATION NUMBER: US/09/825.301
CURRENT FILING DATE: 2001-04-02
NUMBER OF SEQ ID NOS: 77
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-825-301-4

Query Match 65.5%; Score 36; DB 5; Length 292;
Best Local Similarity 70.0%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPILE 12
DB 259 SASIVSPLE 268

RESULT 5
US-09-439-313-532
Sequence 532, Application US/09439313
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439.313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 532
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-439-313-532

Query Match 65.5%; Score 36; DB 5; Length 292;
Best Local Similarity 70.0%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPILE 12
DB 259 SASIVSPLE 268

RESULT 6
US-09-590-583-299
Sequence 299, Application US/09590583
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Reed, Steven G.
APPLICANT: Smith, John M.
APPLICANT: Misher, Linda E.
APPLICANT: Dillon, David C.
APPLICANT: Retter, Marc W.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
FILE REFERENCE: 210121.419C9
CURRENT APPLICATION NUMBER: US/09/590.583
CURRENT FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 324
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 299
LENGTH: 329
TYPE: PRT
ORGANISM: Homo sapien
US-09-590-583-299

Query Match 65.5%; Score 36; DB 5; Length 329;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPLLE 12
:|:|:|:|:|:|
DB 296 SASIVSPLLE 305

RESULT 7
US-09-822-827-376

Sequence 376, Application US/09822827
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.534C1

CURRENT APPLICATION NUMBER: US/09/822,827

CURRENT FILING DATE: 2001-03-28

NUMBER OF SEQ ID NOS: 982

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 376

LENGTH: 329

TYPE: PRT

ORGANISM: Homo sapien

US-09-822-827-376

Query Match 65.5%; Score 36; DB 5; Length 329;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPLLE 12
:|:|:|:|:|:|
DB 296 SASIVSPLLE 305

RESULT 8
US-09-825-301-2

Sequence 2, Application US/09825301
GENERAL INFORMATION:

APPLICANT: Houghton, Raymond L.

APPLICANT: Dillon, David C.

APPLICANT: Molesch, David A.

APPLICANT: Xu, Jiangchun

APPLICANT: Zehentner, Barbara

APPLICANT: Persing, David H.

TITLE OF INVENTION: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION
AND MONITORING OF BREAST CANCER

FILE REFERENCE: 210121.513

CURRENT APPLICATION NUMBER: US/09/825,301

CURRENT FILING DATE: 2001-04-02

NUMBER OF SEQ ID NOS: 77

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 329

TYPE: PRT

ORGANISM: Homo sapien

US-09-825-301-2

Query Match 65.5%; Score 36; DB 5; Length 329;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPLLE 12
:|:|:|:|:|:|
DB 296 SASIVSPLLE 305

RESULT 9
US-09-439-313-376
Sequence 376, Application US/09439313

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, David C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan Louise

APPLICANT: Jiang Yugu

APPLICANT: Reed, Steven G.

APPLICANT: Kalos, Michael

APPLICANT: Fanger, Gary

APPLICANT: Retter, Mark

APPLICANT: Solk, John

APPLICANT: Day, Craig

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C9

CURRENT APPLICATION NUMBER: US/09/439,313

CURRENT FILING DATE: 1999-11-12

NUMBER OF SEQ ID NOS: 575

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 376

LENGTH: 329

TYPE: PRT

ORGANISM: Homo sapien

US-09-439-313-376

Query Match 65.5%; Score 36; DB 5; Length 329;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASVTSPLLE 12
:|:|:|:|:|:|
DB 296 SASIVSPLLE 305

RESULT 10
US-09-739-449-8794

Sequence 8794, Application US/09739449
GENERAL INFORMATION:

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15490)C

CURRENT APPLICATION NUMBER: US/09/739,449

CURRENT FILING DATE: 2000-12-19

PRIOR APPLICATION NUMBER: US 09/514,000

PRIOR FILING DATE: 2000-02-23

NUMBER OF SEQ ID NOS: 13351

SEQ ID NO 8794

LENGTH: 356

TYPE: PRT

ORGANISM: Agrobacterium tumefaciens

US-09-739-449-8794

Query Match 61.8%; Score 34; DB 5; Length 356;
Best Local Similarity 54.5%; Pred. No. 1,1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLNASVTSPLL 11
:|:|:|:|:|:|
DB 285 LNLNPAVTSPLV 295

RESULT 11
US-09-739-449-10537

Sequence 10537, Application US/09739449
GENERAL INFORMATION:

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15490)C

CURRENT APPLICATION NUMBER: US/09/739,449

CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US-09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 10537
LENGTH: 104
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10537

Query Match 58.2%; Score 32; DB 5; Length 104;
Best Local Similarity 70.0%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLNASTVSPL 10
||| |||
DB 31 VLNASTVSPLV 40

RESULT 12
US-09-270-849B-189198
Sequence 189198, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 189198
LENGTH: 164
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-189198

Query Match 58.2%; Score 32; DB 5; Length 164;
Best Local Similarity 60.0%; Pred. No. 1,2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLNASTVSPL 10
|:|:|:|:|:|
DB 127 VVDPSTISPL 136

RESULT 13
US-09-270-849B-194755
Sequence 194755, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 194755
LENGTH: 341
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-194755

Query Match 58.2%; Score 32; DB 5; Length 341;
Best Local Similarity 60.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLNASTVSPL 10
|:|:|:|:|:|
DB 124 VVDPSTISPL 133

RESULT 14
US-09-232-773A-12
Sequence 12, Application US/09232773A
GENERAL INFORMATION:
APPLICANT: Kendall, Richard L.
APPLICANT: Thomas, Kenneth A., Jr.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
FILE REFERENCE: 18888DB
CURRENT APPLICATION NUMBER: US/09/232,773A
CURRENT FILING DATE: 1999-07-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 661
TYPE: PRT
ORGANISM: Homo sapien
US-09-232-773A-12

Query Match 58.2%; Score 32; DB 5; Length 661;
Best Local Similarity 60.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLNASTVSPL 10
|||:|:|:|
DB 223 VLNCTATITPL 232

RESULT 15
US-09-232-773A-6
Sequence 6, Application US/09232773A
GENERAL INFORMATION:
APPLICANT: Kendall, Richard L.
APPLICANT: Thomas, Kenneth A., Jr.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
FILE REFERENCE: 18888DB
CURRENT APPLICATION NUMBER: US/09/232,773A
CURRENT FILING DATE: 1999-07-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 687
TYPE: PRT
ORGANISM: Homo sapien
US-09-232-773A-6

Query Match 58.2%; Score 32; DB 5; Length 687;
Best Local Similarity 60.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

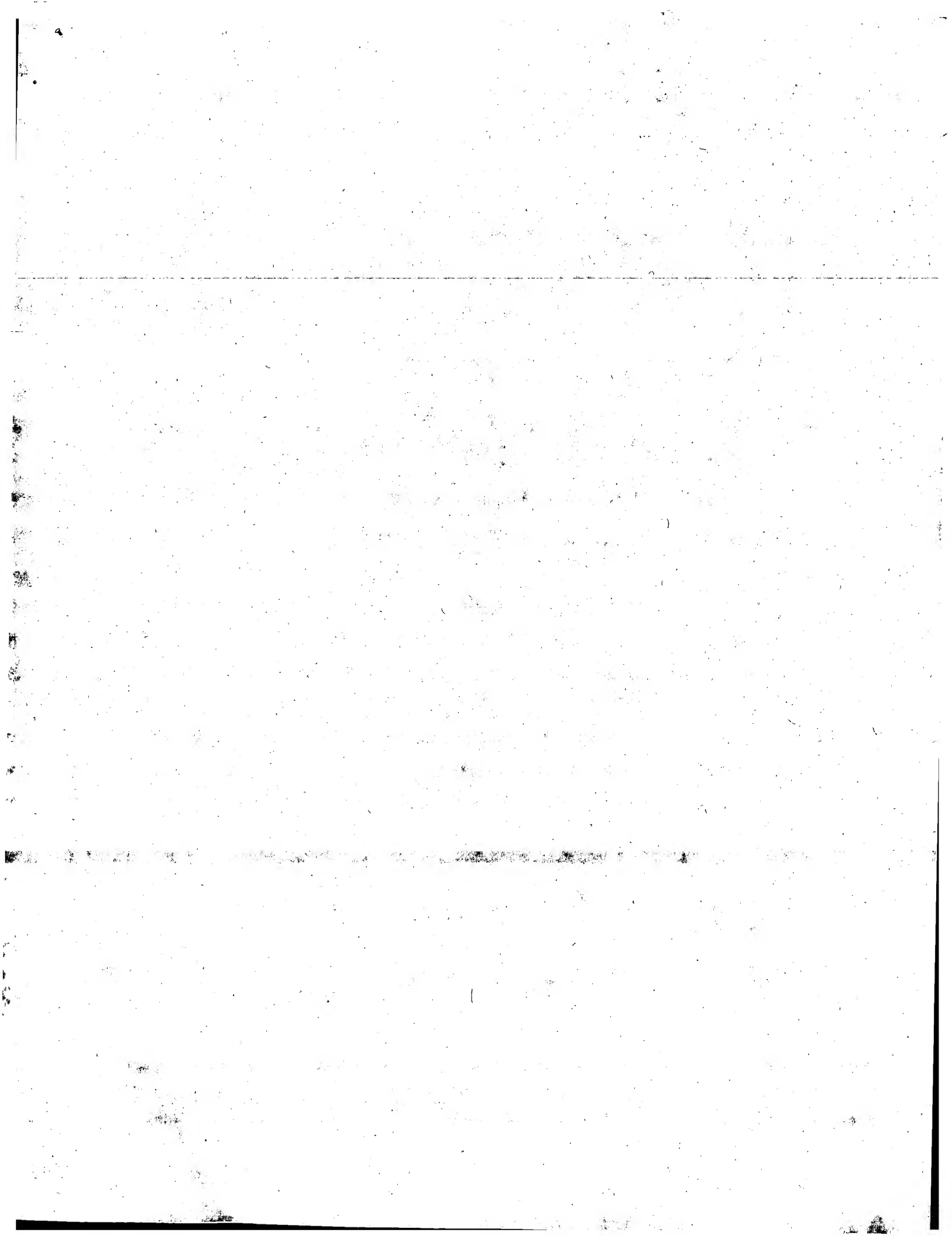
QY 1 VLNASTVSPL 10
|||:|:|:|
DB 249 VLNCTATITPL 258

Search completed: June 4, 2001, 12:21:28
Job time: 590 sec

Tue Jun 5 07:08:19 2001

us-09-284-107-21.rapn

Page 5



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:20:01 ; Search time 260.2 Seconds
(without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-22

Perfect score: 61
Sequence: 1 GNLVTLSCETKL 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents, AA Main: *
1: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US082.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US081.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US080.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	58	23	US-60-160-209-3193
2	61	100.0	96	23	US-60-160-209-4252
3	61	100.0	102	23	US-60-160-203-4414
4	61	100.0	261	16	US-09-245-764-7
5	61	100.0	374	21	US-09-702-021-14
6	61	100.0	399	1	PCT-US00-05882-1128
7	48	78.7	58	21	US-09-702-021-10
8	48	78.7	75	23	US-60-170-373-3220
9	48	78.7	75	23	US-60-170-374-2665
10	48	78.7	91	7	US-08-332-562-94

11	48	78.7	91	8	US-08-462-973-94	Sequence 94, Appl
12	48	78.7	177	23	US-60-185-361-731	Sequence 731, Appl
13	48	78.7	177	23	US-60-185-362-692	Sequence 692, Appl
14	48	78.7	278	21	US-09-702-021-6	Sequence 6, Appl
15	48	78.7	305	21	US-09-702-021-34	Sequence 34, Appl
16	48	78.7	343	21	US-09-702-021-5	Sequence 5, Appl
17	48	78.7	370	21	US-09-702-021-3	Sequence 3, Appl
18	48	78.7	370	21	US-09-702-021-30	Sequence 30, Appl
19	48	78.7	370	21	US-09-702-021-31	Sequence 31, Appl
20	48	78.7	370	21	US-09-702-021-32	Sequence 32, Appl
21	48	78.7	370	21	US-09-702-021-33	Sequence 33, Appl
22	48	78.7	410	3	US-07-861-895-25	Sequence 25, Appl
23	47	77.0	59	23	US-60-170-374-2541	Sequence 2541, Ap
24	47	77.0	96	23	US-60-194-508-3133	Sequence 3133, Ap
25	46	75.4	59	23	US-60-170-374-2665	Sequence 2665, Ap
26	43	70.5	83	23	US-60-195-053-1821	Sequence 1821, Ap
27	43	70.5	181	1	PCT-US01-01326-137	Sequence 137, Appl
28	43	70.5	181	1	PCT-US01-01349-834	Sequence 834, Appl
29	43	70.5	217	1	PCT-US01-01349-664	Sequence 664, Appl
30	43	70.5	359	17	US-09-380-138-45	Sequence 45, Appl
31	42	68.9	358	16	US-09-206-647-1	Sequence 1, Appl
32	42	68.9	359	1	PCT-US00-07285-116	Sequence 116, Appl
33	42	68.9	359	1	PCT-US00-29363-134	Sequence 134, Appl
34	42	68.9	359	19	US-09-528-003-116	Sequence 116, Appl
35	40	65.6	123	23	US-60-173-469-1595	Sequence 1595, Ap
36	39	63.9	59	23	US-60-178-307-2040	Sequence 2040, Ap
37	39	63.9	95	21	US-09-731-872-364	Sequence 364, Appl
38	39	63.9	95	23	US-60-147-499-5383	Sequence 5383, Ap
39	39	63.9	95	23	US-60-169-629-676	Sequence 676, Appl
40	39	63.9	95	23	US-60-187-470-676	Sequence 676, Appl
41	39	63.9	95	23	US-60-197-873-18502	Sequence 18502, A
42	39	63.9	147	1	PCT-US01-01326-158	Sequence 158, Appl
43	39	63.9	147	1	PCT-US01-01326-217	Sequence 217, Appl
44	39	63.9	147	1	PCT-US01-01349-587	Sequence 587, Appl
45	39	63.9	147	1	PCT-US01-01349-785	Sequence 785, Appl

ALIGNMENTS

RESULT 1.
US-60-160-209-3193
GENERAL INFORMATION: Application US/60160209
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
FILE REFERENCE: C1000113
CURRENT APPLICATION NUMBER: US/60/160,209
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FASTSEQ for Windows, Version 4.0
SEQ ID NO 3193
LENGTH: 58
TYPE: PRT
ORGANISM: HUMAN
US-60-160-209-3193

Query Match 100.0% ; Score 61; DB 23; Length 58;
Best Local Similarity 100.0% ; Pred. No. 0.0021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
|||||||
Db 1 GNLVTLSCETKL 12

RESULT 2
US-60-160-209-4252
Sequence 4252, Application US/60160209
GENERAL INFORMATION:

APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
TITLE OF INVENTION: THERMOF
FILE REFERENCE: CL000113
CURRENT APPLICATION NUMBER: US/60/160,209
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4252
LENGTH: 96
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)-(96)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-209-4252

Query Match
Best Local Similarity 100.0%; Score 61; DB 23; Length 96;
Pred. No. 0.0036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
|||||
DB 19 GNLVTLSCETKL 30

RESULT 3
US-60-160-203-4414
Sequence 4414, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
USES THEREOF
FILE REFERENCE: CL000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4414
LENGTH: 102
TYPE: PRT
ORGANISM: HUMAN
US-60-160-203-4414

Query Match
Best Local Similarity 100.0%; Score 61; DB 23; Length 102;
Pred. No. 0.0038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
|||||
DB 27 GNLVTLSCETKL 38

RESULT 4
US-09-245-764-7
Sequence 7, Application US/09245764
GENERAL INFORMATION:
APPLICANT: HOGARTH, P. Mark
APPLICANT: POWELL, MAREE S.
APPLICANT: MCKENZIE, IAN F.C.
APPLICANT: MAXWELL, KELLY F.
APPLICANT: GARRETT, THOMAS P.J.
APPLICANT: EPA, VIDANA
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES AND MODELS OF FC RECEPTORS
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 4102-4
CURRENT APPLICATION NUMBER: US/09/245,764
CURRENT FILING DATE: 1999-02-05

EARLIER APPLICATION NUMBER: 60/099,994
EARLIER FILING DATE: 1998-09-11
EARLIER APPLICATION NUMBER: 60/073,972
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 261
TYPE: PRT
ORGANISM: Homo sapiens
US-09-245-764-7

Query Match
Best Local Similarity 100.0%; Score 61; DB 16; Length 261;
Pred. No. 0.0097;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
|||||
DB 186 GNLVTLSCETKL 197

RESULT 5
US-09-702-021-14
Sequence 14, Application US/09702021
GENERAL INFORMATION:
APPLICANT: C. Fraser
TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF
FILE REFERENCE: 7853-217
CURRENT APPLICATION NUMBER: US/09/702,021
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 374
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-021-14

Query Match
Best Local Similarity 100.0%; Score 61; DB 21; Length 374;
Pred. No. 0.014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
|||||
DB 205 GNLVTLSCETKL 216

RESULT 6
PCT-US00-05882-1128
Sequence 1128, Application PC/TUS0005882
GENERAL INFORMATION:
APPLICANT: CRAIG ROSEN,
APPLICANT: STEVE RUBEN,
TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
FILE REFERENCE: PA106PCT
CURRENT APPLICATION NUMBER: PCT/US00/05882
CURRENT FILING DATE: 2000-03-08
EARLIER APPLICATION NUMBER: 60/124,270
EARLIER FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1128
LENGTH: 399
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (208)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE

Query Match 78.7%; Score 48; DB 7; Length 91;
Best Local Similarity 75.0%; Pred. No. 0.66;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
DB 17 GSLVTLNCETKL 28

RESULT 11

US-08-462-973-94
Sequence 94, Application US/08462973
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONDONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,973
CLASSIFICATION: 435
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/332,562
FILING DATE: 31-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM8232/94
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU87/00159
FILING DATE: 29-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH6166
FILING DATE: 26-MAY-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-462-973-94

Query Match 78.7%; Score 48; DB 8; Length 91;
Best Local Similarity 75.0%; Pred. No. 0.66;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
DB 17 GSLVTLNCETKL 28

RESULT 12

US-60-185-361-731
Sequence 731, Application US/60185361
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
FILE REFERENCE: CL000291
CURRENT APPLICATION NUMBER: US/60/185,361
CURRENT FILING DATE: 2000-02-28
NUMBER OF SEQ ID NOS: 968
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 731
LENGTH: 177
TYPE: PRT
ORGANISM: HUMAN
US-60-185-361-731

Query Match 78.7%; Score 48; DB 23; Length 177;
Best Local Similarity 75.0%; Pred. No. 1.3;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
DB 2 GSLVTLNCOTKL 13

RESULT 13

US-60-185-362-692
Sequence 692, Application US/60185362
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
FILE REFERENCE: CL000285
CURRENT APPLICATION NUMBER: US/60/185,362
CURRENT FILING DATE: 2000-02-28
NUMBER OF SEQ ID NOS: 836
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 692
LENGTH: 177
TYPE: PRT
ORGANISM: HUMAN
US-60-185-362-692

Query Match 78.7%; Score 48; DB 23; Length 177;
Best Local Similarity 75.0%; Pred. No. 1.3;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
DB 2 GSLVTLNCOTKL 13

RESULT 14

US-09-702-021-6
Sequence 6, Application US/09702021
GENERAL INFORMATION:
APPLICANT: C. Fraser
TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF
FILE REFERENCE: 7853-217
CURRENT APPLICATION NUMBER: US/09/702,021
CURRENT FILING DATE: 2000-10-30

NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 6
LENGTH: 278
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-021-6

Query Match 78.7%; Score 48; DB 21; Length 278;
Best Local Similarity 75.0%; Pred. No. 2;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
|:|||||:||||
DB 93 GSLVTLRCQTKL 104

RESULT 15

US-09-702-021-34
Sequence 34, Application US/09702021
GENERAL INFORMATION:
APPLICANT: C. Fraser
TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF
FILE REFERENCE: 7853-217
CURRENT APPLICATION NUMBER: US/09/702,021
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 34
LENGTH: 305
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-021-34

Query Match 78.7%; Score 48; DB 21; Length 305;
Best Local Similarity 75.0%; Pred. No. 2.2;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
|:|||||:||||
DB 120 GSLVTLRCQTKL 131

Search completed: June 4, 2001, 12:20:01
Job time: 514 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:21:28 ; Search time 84.28 seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-22

Sequence: 1 GNWLTLSCETKL 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA.New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US01_11988-133
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	12	US-09-284-107-22	Sequence 22, Appl
2	42	68.9	359	PCT-US01-11988-133	Sequence 133, App
3	42	68.9	359	US-09-833-245-133	Sequence 133, App
4	38	62.3	257	US-09-270-849B-180519	Sequence 180519,
5	38	62.3	544	US-09-565-522B-12	Sequence 12, Appl
6	37	60.7	287	US-09-723-716A-43	Sequence 43, Appl
7	37	60.7	287	US-09-723-716A-68	Sequence 68, Appl
8	37	60.7	305	US-09-723-716A-70	Sequence 70, Appl
9	35	57.4	57	US-09-514-009-46	Sequence 46, Appl
10	35	57.4	107	US-09-514-009-36	Sequence 36, Appl
11	35	57.4	205	US-09-514-009-34	Sequence 34, Appl
12	35	57.4	239	US-09-514-009-33	Sequence 33, Appl
13	35	57.4	479	US-09-184-467-20	Sequence 20, Appl
14	35	57.4	479	US-09-184-467-24	Sequence 24, Appl
15	35	57.4	507	US-09-184-467-16	Sequence 16, Appl
16	35	57.4	507	US-09-184-467-28	Sequence 28, Appl
17	34.5	56.6	571	US-60-248-505-847	Sequence 847, Appl
18	34.5	56.6	571	US-60-248-505-1220	Sequence 1220, Ap
19	34	55.7	158	US-09-345-236B-22	Sequence 22, Appl
20	34	55.7	174	US-09-345-236B-22	Sequence 22, Appl
21	34	55.7	357	US-09-739-449-12563	Sequence 12563, A
22	34	55.7	495	PCT-US01-11988-1679	Sequence 1679, Ap
23	34	55.7	495	PCT-US01-11988-1680	Sequence 1680, Ap
24	34	55.7	495	US-09-833-245-1679	Sequence 1679, Ap
25	34	55.7	495	US-09-833-245-1680	Sequence 1680, Ap
26	34	55.7	1778	PCT-US01-04098A-3126	Sequence 3126, Ap
27	34	55.7	2099	PCT-US01-04098A-1158	Sequence 1158, Ap

28	33	54.1	61	US-08-849-406-38	Sequence 38, Appl
29	33	54.1	103	PCT-US01-11988-208	Sequence 208, App
30	33	54.1	103	US-09-833-245-208	Sequence 208, App
31	33	54.1	249	US-09-832-312-9	Sequence 9, Appl
32	33	54.1	268	PCT-US01-10462-13	Sequence 13, Appl
33	33	54.1	319	US-09-832-312-5	Sequence 5, Appl
34	33	54.1	339	US-09-832-312-3	Sequence 3, Appl
35	33	54.1	339	US-09-832-312-34	Sequence 34, Appl
36	33	54.1	339	US-09-832-312-36	Sequence 36, Appl
37	33	54.1	339	US-09-832-312-38	Sequence 38, Appl
38	33	54.1	339	US-09-832-312-40	Sequence 40, Appl
39	33	54.1	444	PCT-US00-35017A-1153	Sequence 1153, Ap
40	33	54.1	444	PCT-US00-35017A-1264	Sequence 1264, Ap
41	33	54.1	526	US-09-828-825-4	Sequence 4, Appl
42	33	54.1	527	US-09-828-825-2	Sequence 2, Appl
43	33	54.1	537	US-09-828-825-8	Sequence 8, Appl
44	33	54.1	537	US-09-828-825-16	Sequence 16, Appl
45	33	54.1	542	US-09-828-825-6	Sequence 6, Appl

ALIGNMENTS

```
RESULT 1
US-09-284-107-22
Sequence 22, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
APPLICANT: de Krulff, Cornelis Adriaan John
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: AND PROTEINS HAVING SPECIFIC AFFINITY FOR A TARGET
FILE REFERENCE: 313632000600
CURRENT APPLICATION NUMBER: US/09/284,107
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
PRIOR FILING DATE: 1996-10-08
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 22
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
US-09-284-107-22

Query Match 100.0%; Score 61; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 96-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNWLTLSCETKL 12
DB 1 GNWLTLSCETKL 12

RESULT 2
PCT-US01-11988-133
Sequence 133, Application PCT/US0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
```

PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 133
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-133

Query Match 68.9%; Score 42; DB 1; Length 359;
Best Local Similarity 66.7%; Pred. No. 6;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
| :||||:|
Db 185 GGPMTLSCOTKL 196

RESULT 3
US-09-833-245-133
Sequence 133, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCR
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 133
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-133

Query Match 68.9%; Score 42; DB 5; Length 359;
Best Local Similarity 66.7%; Pred. No. 6;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
| :||||:|
Db 185 GGPMTLSCOTKL 196

RESULT 4
US-09-270-849B-180519
Sequence 180519, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 180519
LENGTH: 257
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-180519

Query Match 62.3%; Score 38; DB 5; Length 257;

Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 LVLSCETKL 12
| :||||:|
Db 10 LFLFSCETKL 19

RESULT 5
US-09-565-522B-12
Sequence 12, Application US/09565522B
GENERAL INFORMATION:
APPLICANT: Boyle, Bryan J
APPLICANT: Mize, Nancy K
APPLICANT: Alterburn, Matthew C
APPLICANT: Tang, Y. Tom
APPLICANT: Tillinghast, John
APPLICANT: Slinku, Ankura
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Methods and Materials Relating to Novel Interleukin-1 Receptor
FILE REFERENCE: HYS-8
CURRENT APPLICATION NUMBER: US/09/565,522B
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: US/09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US/09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 544
TYPE: PRT
ORGANISM: Homo sapiens
US-09-565-522B-12

Query Match 62.3%; Score 38; DB 5; Length 544;
Best Local Similarity 54.3%; Pred. No. 47;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 11
| :||||:|
Db 78 GDVTVACEAK 88

RESULT 6
US-09-723-716A-43
Sequence 43, Application US/09723716A
GENERAL INFORMATION:
APPLICANT: JAKOBSEN, Bent Karsten
APPLICANT: BELL, John Irving
APPLICANT: GAO, George Fu
APPLICANT: WILLCOX, Benjamin Ernest
APPLICANT: BOULTER, Jonathan Michael
TITLE OF INVENTION: Soluble T Cell Receptor
FILE REFERENCE: 102286.409DIV
CURRENT APPLICATION NUMBER: US/09/723,716A
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US/09/335,087
PRIOR FILING DATE: 1999-06-17
PRIOR APPLICATION NUMBER: PCT/GB99/01588
PRIOR FILING DATE: 1999-05-19
PRIOR APPLICATION NUMBER: GB 9810759.2
PRIOR FILING DATE: 1998-05-19
PRIOR APPLICATION NUMBER: GB 9821129.5
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 2.1
SEQ ID NO 43
LENGTH: 287
TYPE: PRT

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Predicted amino acid sequence of the soluble
OTHER INFORMATION: HLA-A2/flu matrix restricted beta chain from JM22,
OTHER INFORMATION: as fused to the "leucine zipper" domain of c-fos.
OTHER INFORMATION: (Figure 11).
US-09-723-716A-43

Query Match 60.7%; Score 37; DB 5; Length 287;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
| | | | | | | |
DB 18 GQNVTLSCQNL 29

RESULT 7
US-09-723-716A-68
Sequence 68, Application US/09723716A
GENERAL INFORMATION:
APPLICANT: JAKOBSEN, Bent Karsten
APPLICANT: BELT, John Irving
APPLICANT: GAO, George Fu
APPLICANT: WILCOX, Benjamin Ernest
APPLICANT: BOULTER, Jonathan Michael
TITLE OF INVENTION: Soluble T Cell Receptor
FILE REFERENCE: 102286.409D1V
CURRENT APPLICATION NUMBER: US/09/723,716A
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 09/335,087
PRIOR FILING DATE: 1999-06-17
PRIOR APPLICATION NUMBER: PCT/GB99/01588
PRIOR FILING DATE: 1999-05-19
PRIOR APPLICATION NUMBER: GB 9810759.2
PRIOR FILING DATE: 1998-05-19
PRIOR APPLICATION NUMBER: GB 9821129.5
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 2.1
SEQ ID NO 68
LENGTH: 287
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Amino acid
OTHER INFORMATION: sequence of human HLA-A2/flu matrix peptide
OTHER INFORMATION: restricted JM22 TCR beta chain fused to c-fos
OTHER INFORMATION: leucine zipper domain.
US-09-723-716A-68

Query Match 60.7%; Score 37; DB 5; Length 287;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
| | | | | | | |
DB 18 GQNVTLSCQNL 29

RESULT 8
US-09-723-716A-70
Sequence 70, Application US/09723716A
GENERAL INFORMATION:
APPLICANT: JAKOBSEN, Bent Karsten
APPLICANT: BELT, John Irving
APPLICANT: GAO, George Fu
APPLICANT: WILCOX, Benjamin Ernest
APPLICANT: BOULTER, Jonathan Michael
TITLE OF INVENTION: Soluble T Cell Receptor
FILE REFERENCE: 102286.409D1V

CURRENT APPLICATION NUMBER: US/09/723,716A
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 09/335,087
PRIOR FILING DATE: 1999-06-17
PRIOR APPLICATION NUMBER: PCT/GB99/01588
PRIOR FILING DATE: 1999-05-19
PRIOR APPLICATION NUMBER: GB 9810759.2
PRIOR FILING DATE: 1998-05-19
PRIOR APPLICATION NUMBER: GB 9821129.5
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 2.1
SEQ ID NO 70
LENGTH: 305
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Amino acid
OTHER INFORMATION: sequence of human HLA-A2/flu matrix peptide
OTHER INFORMATION: restricted JM22 TCR beta chain fused to c-fos
OTHER INFORMATION: leucine zipper domain and BirA biotinylation tag.
US-09-723-716A-70

Query Match 60.7%; Score 37; DB 5; Length 305;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
| | | | | | | |
DB 18 GQNVTLSCQNL 29

RESULT 9
US-09-514-009-46
Sequence 46, Application US/09514009
GENERAL INFORMATION:
APPLICANT: Fraser, Christopher C.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 07334-251001
CURRENT APPLICATION NUMBER: US/09/514,009
CURRENT FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/259,387
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 46
LENGTH: 57
TYPE: PRT
ORGANISM: Mus musculus
US-09-514-009-46

Query Match 57.4%; Score 35; DB 5; Length 57;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 11
| | | | | | | |
DB 1 GQNVTLSCQNL 11

RESULT 10
US-09-514-009-36
Sequence 36, Application US/09514009
GENERAL INFORMATION:
APPLICANT: Fraser, Christopher C.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 07334-251001
CURRENT APPLICATION NUMBER: US/09/514,009
CURRENT FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/259,387
PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 107
TYPE: PRT
ORGANISM: Mus musculus
US-09-514-009-36

Query Match 57.4%; Score 35; DB 5; Length 107;
Best Local Similarity 63.6%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
| : ||| |
Db 25 GQNVSLSCSTK 35

RESULT 11
US-09-514-009-34
Sequence 34, Application US/09514009
GENERAL INFORMATION:
APPLICANT: Fraser, Christopher C.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 07334-251001
CURRENT APPLICATION NUMBER: US/09/514,009
CURRENT FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/259,387
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 34
LENGTH: 205
TYPE: PRT
ORGANISM: Mus musculus
US-09-514-009-34

Query Match 57.4%; Score 35; DB 5; Length 205;
Best Local Similarity 63.6%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
| : ||| |
Db 25 GQNVSLSCSTK 35

RESULT 12
US-09-514-009-33
Sequence 33, Application US/09514009
GENERAL INFORMATION:
APPLICANT: Fraser, Christopher C.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 07334-251001
CURRENT APPLICATION NUMBER: US/09/514,009
CURRENT FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/259,387
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 239
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(34)
US-09-514-009-33

Query Match 57.4%; Score 35; DB 5; Length 239;
Best Local Similarity 63.6%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
| : ||| |
Db 59 GQNVSLSCSTK 69

RESULT 13
US-09-184-467-20
Sequence 20, Application US/09184467
GENERAL INFORMATION:
APPLICANT: Hasel, Karl W.
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED
THEREOF, AND THERAPEUTIC AND PROPHYLACTIC USES
THEREOF
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/184,467
FILING DATE: 02-Nov-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/530,146
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41190/JFW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOPBU
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-184-467-20

Query Match 57.4%; Score 35; DB 5; Length 479;
Best Local Similarity 54.5%; Pred. No. 1,5e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
| : ||| |
Db 380 GNTILPCRIK 390

RESULT 14
US-09-184-467-24
Sequence 24, Application US/09184467
GENERAL INFORMATION:
APPLICANT: Hasel, Karl W.
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED
THEREOF, AND THERAPEUTIC AND PROPHYLACTIC USES
THEREOF
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.24

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/184,467
FILING DATE: 02-Nov-1998
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/530,146
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41190/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOPUI

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-184-467-24

Query Match 57.4%; Score 35; DB 5; Length 479;
Best Local Similarity 54.5%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11
||:|||||
Db 380 GNTITLPCRIK 390

RESULT 15
US-09-184-467-16
Sequence 16, Application US/09184467
GENERAL INFORMATION:
APPLICANT: Hasel, Karl W.
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED
THEREOF, AND THERAPEUTIC AND PROPHYLACTIC USES
THEREOF
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza,
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.24

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/184,467
FILING DATE: 02-Nov-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/530,146

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41190/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOPUI

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-184-467-16

Query Match 57.4%; Score 35; DB 5; Length 507;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11
||:|||||
Db 408 GNTITLPCRIK 418

Search completed: June 4, 2001, 12:21:28
Job time: 590 sec

Tue Jun 5 07:08:32 2001

us-09-284-107-22.rapn

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:20:01 ; Search time 260.2 Seconds
(without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-23
Perfect score: 61
Sequence: 1 LLORPGQLYFS 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues
Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/US06.COMB.pep:*
- 2: /cgn2_6/ptodata/2/paa/US07.COMB.pep:*
- 3: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 4: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 5: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 6: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 7: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 8: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 9: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 10: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 11: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 12: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 13: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 14: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 15: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 16: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 17: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 18: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 19: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 20: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 21: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 22: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
- 23: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	58	US-60-160-209-3193	Sequence 3193, Ap
2	61	100.0	96	US-60-160-209-4252	Sequence 4252, Ap
3	61	100.0	102	US-60-160-203-4414	Sequence 4414, Ap
4	61	100.0	374	US-09-702-021-14	Sequence 14, Appl
5	61	100.0	399	PCT-US00-05882-1128	Sequence 1128, Ap
6	56	91.8	91	US-08-332-562-94	Sequence 94, Appl
7	56	91.8	91	US-08-462-973-94	Sequence 94, Appl
8	56	91.8	410	US-07-861-895-25	Sequence 25, Appl
9	55	90.2	261	US-09-245-764-7	Sequence 7, Appl
10	39	63.9	373	US-60-146-394-1032	Sequence 1032, Ap

	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45		
	39	39	39	39	39	39	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	37	37	37	37	37	37	37	37	36	36	36	36
	63.9	63.9	63.9	63.9	63.9	63.9	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	62.3	60.7	60.7	60.7	60.7	60.7	60.7	60.7	60.7	60.7	59.0	59.0	59.0	59.0
	535	611	612	612	612	612	787	787	787	787	787	787	787	787	787	787	787	787	787	787	787	787	787	787	787	350	350	350	350	350	350	350	350	350	350	350	350
	US-09-489-039A-13963	US-60-146-055-774	US-60-167-217-452	US-60-173-464-405	US-60-191-637-477	US-60-191-681-387	US-09-252-991A-17591	US-60-229-515-758	US-08-720-484-4	US-08-897-798-6	US-08-953-823-4	US-09-398-239-4	US-09-711-111-6	US-08-720-484-2	US-08-897-798-7	US-08-953-823-2	US-09-060-939A-16	US-09-060-939A-17	US-09-293-505-15	US-09-293-505-17	US-09-398-239-2	US-09-711-111-7	US-09-060-939A-15	US-09-293-505-15	US-09-489-039A-9855	US-09-328-352-6649	US-09-252-691-7682	US-09-252-691-7682	US-60-161-932-2494	US-60-173-464-18875	US-60-191-681-18143	US-09-252-991A-28773	US-60-191-681-18143	US-60-191-681-18143	US-60-191-681-18143	US-60-191-681-18143	US-60-191-681-18143
	Sequence 13963, A	Sequence 774, App	Sequence 452, App	Sequence 405, App	Sequence 477, App	Sequence 387, App	Sequence 17591, A	Sequence 758, App	Sequence 4, Appl	Sequence 6, Appl	Sequence 4, Appl	Sequence 4, Appl	Sequence 6, Appl	Sequence 2, Appl	Sequence 7, Appl	Sequence 2, Appl	Sequence 16, Appl	Sequence 17, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 17, Appl	Sequence 7, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 9855, App	Sequence 6649, App	Sequence 7682, App	Sequence 7682, App	Sequence 2494, App	Sequence 18875, A	Sequence 22977, A	Sequence 18143, A	Sequence 28773, A	Sequence 26977, A	Sequence 21586, A		

ALIGNMENTS

RESULT 1
US-60-160-209-3193
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
FILE REFERENCE: CLO00113
CURRENT APPLICATION NUMBER: US/60/160,209
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 3193
LENGTH: 58
TYPE: PRT
ORGANISM: HUMAN
US-60-160-209-3193

Query Match: 100.0%; Score 61; DB 23; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGQLYFS 12
|||||
Db 13 LLORPGQLYFS 24

RESULT 2
US-60-160-209-4252
Sequence 4252, Application US/60160209
GENERAL INFORMATION:

APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: C1000113
CURRENT APPLICATION NUMBER: US/60/160,209
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4252
LENGTH: 96
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(96)
OTHER INFORMATION: Xaa - Any Amino Acid
US-60-160-209-4252

Query Match
Best Local Similarity 100.0%; Score 61; DB 23; Length 96;
Matches 12; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGLQLYFS 12
DB 31 LLORPGLQLYFS 42

RESULT 3
US-60-160-203-4414
Sequence 4414, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
USE THEREOF
FILE REFERENCE: C1000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4414
LENGTH: 102
TYPE: PRT
ORGANISM: HUMAN
US-60-160-203-4414

Query Match
Best Local Similarity 100.0%; Score 61; DB 23; Length 102;
Matches 12; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGLQLYFS 12
DB 39 LLORPGLQLYFS 50

RESULT 4
US-09-702-021-14
Sequence 14, Application US/09702021
GENERAL INFORMATION:
APPLICANT: C. Fraser
TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF
FILE REFERENCE: 7853-217
CURRENT APPLICATION NUMBER: US/09/702,021
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 14
LENGTH: 374
TYPE: PRT
ORGANISM: Homo sapiens

US-09-702-021-14

Query Match
Best Local Similarity 100.0%; Score 61; DB 21; Length 374;
Matches 12; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGLQLYFS 12
DB 217 LLORPGLQLYFS 228

RESULT 5
PCT-US00-05882-1128
Sequence 1128, Application PC/TUS0005882
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
Applicant: Steve Ruben
TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
FILE REFERENCE: P4106PCT
CURRENT APPLICATION NUMBER: PCT/US00/05882
CURRENT FILING DATE: 2000-03-08
EARLIER APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1128
LENGTH: 399
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (208)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (349)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1128

Query Match
Best Local Similarity 100.0%; Score 61; DB 1; Length 399;
Matches 12; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGLQLYFS 12
DB 242 LLORPGLQLYFS 253

RESULT 6
US-08-332-562-94
Sequence 94, Application US/08332562
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONDONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: PC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562-94

Query Match 91.8%; Score 56; DB 7; Length 91;
Best Local Similarity 91.7%; Pred. No. 0.013;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIORPGLQLYFS 12
DB 29 LIORPGLQLYFS 40

RESULT 7
US-08-462-973-94
Sequence 94, Application US/08462973
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,973
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/332,562
FILING DATE: 31-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM8232/94
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU87/000159

FILING DATE: 29-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH6166
FILING DATE: 26-MAY-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-462-973-94

Query Match 91.8%; Score 56; DB 8; Length 91;
Best Local Similarity 91.7%; Pred. No. 0.013;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIORPGLQLYFS 12
DB 29 LIORPGLQLYFS 40

RESULT 8
US-07-861-895-25
Sequence 25, Application US/07861895
GENERAL INFORMATION:
APPLICANT: HOGARTH, Phillip M.
APPLICANT: HULETT, Mark D.
APPLICANT: IERINO, Francesco L.
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: OSMAN, Nalin
TITLE OF INVENTION: HYBRID FC RECEPTOR MOLECULES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/861,895
FILING DATE: 19920624
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00513
FILING DATE: 25-OCT-1990.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7045
FILING DATE: 25-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17227/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids

TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-861-895-25

Query Match: 91.8%; Score 56; DB 3; Length 410;
Best Local Similarity 91.7%; Pred. No. 0.063;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLOPGLQLYFS 12
DB 232 LLOPGLQLYFS 243

RESULT 9
US-09-245-764-7
Sequence 7, Application US/09245764
GENERAL INFORMATION:
APPLICANT: Hogarth, P. Mark
APPLICANT: Powell, Maree S.
APPLICANT: McKenzie, Ian F.C.
APPLICANT: Maxwell, Kelly F.
APPLICANT: Garrett, Thomas P.J.
APPLICANT: Epa, Vidana
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES AND MODELS OF FC RECEPTORS
FILE REFERENCE: 4102-4
CURRENT APPLICATION NUMBER: US/09/245,764
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/099,994
EARLIER FILING DATE: 1998-09-11
EARLIER APPLICATION NUMBER: 60/073,972
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 7
LENGTH: 261
TYPE: PRT
ORGANISM: Homo sapiens
US-09-245-764-7

Query Match: 90.2%; Score 55; DB 16; Length 261;
Best Local Similarity 91.7%; Pred. No. 0.06;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLOPGLQLYFS 12
DB 198 LLOPGLQLYFS 209

RESULT 10
US-60-146-394-1032
Sequence 1032, Application US/60146394
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
FILE REFERENCE: C10000063
CURRENT APPLICATION NUMBER: US/60/146,394
CURRENT FILING DATE: 1999-07-29
NUMBER OF SEQ ID NOS: 1866
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1032
LENGTH: 373
TYPE: PRT
ORGANISM: Drosophila
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(373)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-146-394-1032

Query Match: 63.9%; Score 39; DB 23; Length 373;
Best Local Similarity 77.8%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 QRPGLQLYF 11
DB 354 QRPGLQLYF 362

RESULT 11
US-09-489-039A-13963
Sequence 13963, Application US/09489039A
GENERAL INFORMATION:
APPLICANT: Gary Bleton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13963
LENGTH: 535
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13963

Query Match: 63.9%; Score 39; DB 18; Length 535;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLOPGLQLY 10
DB 116 LLOPGLQLY 125

RESULT 12
US-60-146-055-774
Sequence 774, Application US/60146055
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
FILE REFERENCE: C10000067
CURRENT APPLICATION NUMBER: US/60/146,055
CURRENT FILING DATE: 1999-07-28
NUMBER OF SEQ ID NOS: 938
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 774
LENGTH: 611
TYPE: PRT
ORGANISM: Drosophila
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(611)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-146-055-774

Query Match: 63.9%; Score 39; DB 23; Length 611;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 QRPGLQLYF 11
DB 414 QRPGLQLYF 422

RESULT 13

```

US-60-167-217-452
; Sequence 452, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: LI, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000152
; CURRENT APPLICATION NUMBER: US/60/167,217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 452
; LENGTH: 612
; TYPE: PRF
; ORGANISM: Drosophila
US-60-167-217-452

```

```

Query Match      63.9%; Score 39; DB 23; Length 612;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      3 ORPGLQLYF 11
        :|||
Db      415 ERPDQLYF 423

```

```

RESULT 14
US-60-173-464-405
; Sequence 405, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: LI, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 405
; LENGTH: 612
; TYPE: PRF
; ORGANISM: Drosophila
US-60-173-464-405

```

```

Query Match      63.9%; Score 39; DB 23; Length 612;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      3 ORPGLQLYF 11
        :|||
Db      415 ERPDQLYF 423

```

```

RESULT 15
US-60-191-637-477
; Sequence 477, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 477
; LENGTH: 612
; TYPE: PRF

```

```

; ORGANISM: DROSOPHILA
US-60-191-637-477

```

```

Query Match      63.9%; Score 39; DB 23; Length 612;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      3 ORPGLQLYF 11
        :|||
Db      415 ERPDQLYF 423

```

```

Search completed: June 4, 2001, 12:20:02
Job time: 515 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:21:28 ; Search time 84.28 Seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-23
Perfect score: 61
Sequence: 1 LLORPGQLYFS 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents, AA, New: *
1: /cgn2_6/ptodata/1/paa/PCT_US01-11988-1866.ppt: *
2: /cgn2_6/ptodata/1/paa/US01-11988-1866.ppt: *
3: /cgn2_6/ptodata/1/paa/US01-11988-1866.ppt: *
4: /cgn2_6/ptodata/1/paa/US01-11988-1866.ppt: *
5: /cgn2_6/ptodata/1/paa/US01-11988-1866.ppt: *
6: /cgn2_6/ptodata/1/paa/US01-11988-1866.ppt: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	12	5	US-09-284-107-23
2	34.5	56.6	259	5	US-09-270-849B-188156
3	34	55.7	47	1	PCT-US01-11988-1866
4	34	55.7	47	5	US-09-833-245-1866
5	34	55.7	63	1	PCT-US01-11988-1866
6	34	55.7	63	5	US-09-833-245-1866
7	34	55.7	471	5	US-09-311-626B-8
8	33	54.1	370	5	US-09-270-849B-186298
9	33	54.1	391	5	US-09-739-449-11862
10	33	54.1	393	5	US-09-739-449-9884
11	33	54.1	2910	5	US-09-828-498-2
12	32.5	53.3	291	5	US-09-739-449-11020
13	32	52.5	123	5	US-09-724-475-3
14	32	52.5	124	5	US-09-769-066-21
15	32	52.5	141	5	US-09-270-849B-191061
16	32	52.5	184	5	US-09-739-449-11176
17	32	52.5	197	5	US-09-270-849B-189872
18	32	52.5	261	5	US-09-685-330-8
19	32	52.5	277	5	US-09-739-449-12957
20	32	52.5	370	1	PCT-US01-11988-1481
21	32	52.5	370	5	US-09-11988-1482
22	32	52.5	370	5	US-09-685-330-2
23	32	52.5	370	5	US-09-685-330-6
24	32	52.5	370	5	US-09-823-033-5
25	32	52.5	370	5	US-09-833-245-1481
26	32	52.5	370	5	US-09-833-245-1482
27	32	52.5	386	5	US-09-739-449-13136

28	32	52.5	387	1	PCT-US00-35017A-1125	Sequence 1125, Ap
29	32	52.5	2356	6	US-60-285-697-38	Sequence 38, Appl
30	31	50.8	42	5	US-09-601-600-180	Sequence 180, Appl
31	31	50.8	112	5	US-09-772-120-6	Sequence 149, Appl
32	31	50.8	112	5	US-09-647-468-149	Sequence 150, Appl
33	31	50.8	112	5	US-09-647-468-150	Sequence 150, Appl
34	31	50.8	112	5	US-09-262-724-6	Sequence 6, Appl
35	31	50.8	124	5	US-09-769-066-22	Sequence 22, Appl
36	31	50.8	131	5	US-09-647-468-163	Sequence 163, Appl
37	31	50.8	131	5	US-09-647-468-164	Sequence 164, Appl
38	31	50.8	154	1	PCT-US00-35017A-988	Sequence 988, Appl
39	31	50.8	194	5	US-09-308-823A-197	Sequence 197, Appl
40	31	50.8	204	5	US-09-811-284-195	Sequence 195, Appl
41	31	50.8	290	5	US-09-270-849B-182901	Sequence 182901, Appl
42	31	50.8	315	6	US-60-248-505-1125	Sequence 1125, Ap
43	31	50.8	321	5	US-09-270-849B-180893	Sequence 180893, Appl
44	31	50.8	389	5	US-09-739-449-8238	Sequence 8238, Ap
45	31	50.8	479	5	US-09-270-849B-193923	Sequence 193923, Appl

ALIGNMENTS

RESULT 1
US-09-284-107-23

Sequence 23, Application US/09284107

GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton

TITLE OF INVENTION: de Kruijf, Cornelis Adriaan John
METHODS AND MEANS FOR SELECTING PEPTIDES
SPECIFIC AFFINITY FOR A TARGET

FILE REFERENCE: 313632000600
CURRENT APPLICATION NUMBER: US/09/284,107
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07

PRIOR APPLICATION NUMBER: EP 96202791.8
PRIOR FILING DATE: 1996-10-08
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 23
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
domain of CD64

US-09-284-107-23

Query Match 100.0%; Score 61; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGQLYFS 12
DB 1 LLORPGQLYFS 12

US-09-270-849B-188156
Sequence 188156, Application US/09270849B

GENERAL INFORMATION:
APPLICANT: Swimmer et al.

TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 188156
LENGTH: 259

TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-186156

Query Match
Best Local Similarity 56.6%; Score 34.5; DB 5; Length 259;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 2 LORPGLQLYFS 12
||| | | | | |
Db 178 LORPG-ECYFS 187

RESULT 3
PCT-US01-11988-1866
Sequence 1866, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1866
LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-1866

Query Match
Best Local Similarity 55.7%; Score 34; DB 1; Length 47;
Matches 10; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 LORPGL-----QLYFS 12
| | | | | | | | | |
Db 3 LORPGLVIALPPOLFS 20

RESULT 4
US-09-833-245-1866
Sequence 1866, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1866
LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-1866

Query Match
Best Local Similarity 55.7%; Score 34; DB 5; Length 47;
Matches 10; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 LORPGL-----QLYFS 12
| | | | | | | | | |
Db 3 LORPGLVIALPPOLFS 20

RESULT 5
PCT-US01-11988-1864
Sequence 1864, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1864
LENGTH: 63
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (37)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-11988-1864

Query Match
Best Local Similarity 55.7%; Score 34; DB 1; Length 63;
Matches 10; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 LORPGL-----QLYFS 12
| | | | | | | | | |
Db 3 LORPGLVIALPPOLFS 20

RESULT 6
US-09-833-245-1864
Sequence 1864, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1864
LENGTH: 63
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (37)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-1864

Query Match 55.7%; Score 34; DB 5; Length 63;
Best Local Similarity 55.6%; Pred. No. 40;
Matches 10; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 LQRPGLV-----OLYPS 12
| | | | | | | | | | | | | | | | | |
Db 3 LQRPGLVLLPPLSFS 20

RESULT 7
US-09-311-626B-8

; Sequence 8, Application US/09311626B
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schmorl, Kirk
; APPLICANT: Andersen, Lene Nonboe
; APPLICANT: Schulten, Martin
; APPLICANT: Outtrup, Helle
; TITLE OF INVENTION: Novel Rhamnogalacturonan Hydrolases
; FILE REFERENCE: 5572.204-US
; CURRENT APPLICATION NUMBER: US/09/311,626B
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 0608/98
; PRIOR FILING DATE: 1998-05-01
; PRIOR APPLICATION NUMBER: 60/084,358
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-311-626B-8

Query Match 55.7%; Score 34; DB 5; Length 471;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LQRPGLV 10
| | | | | | | | | | | | | | | | | |
Db 263 LLSRPGLEVF 272

RESULT 8
US-09-270-849B-186298

; Sequence 186298, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 186298
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-186298

Query Match 54.1%; Score 33; DB 5; Length 370;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLQLTYS 12
| | | | | | | | | | | | | | | | | |
Db 243 RPNLXLYS 251

RESULT 9
US-09-739-449-11862

; Sequence 11862, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 11862
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11862

Query Match 54.1%; Score 33; DB 5; Length 391;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LQRPGLV 10
| | | | | | | | | | | | | | | | | |
Db 227 LTRPGIALY 235

RESULT 10
US-09-739-449-9884

; Sequence 9884, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9884
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9884

Query Match 54.1%; Score 33; DB 5; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PGLQLV 10
| | | | | | | | | | | | | | | | | |
Db 308 PGLQLV 313

RESULT 11
US-09-828-498-2

; Sequence 2, Application US/09828498
; GENERAL INFORMATION:
; APPLICANT: XIANG, JINHUA
; APPLICANT: WUNSCHMANN, SABINA
; APPLICANT: SCHMIDT, WARREN
; APPLICANT: STAPLETON, JACK T.
; TITLE OF INVENTION: FULL-LENGTH GB VIRUS C (HEPATITIS G VIRUS) RNA TRANSCRIPTS
; FILE REFERENCE: IOMA:0300US
; CURRENT APPLICATION NUMBER: US/09/828,498
; CURRENT FILING DATE: 2001-04-05

PRIOR APPLICATION NUMBER: 60/253,390
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/195,597
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 2910
TYPE: PRT
ORGANISM: Hepatitis G virus
US-09-828-498-2

Query Match 54.1%; Score 33; DB 5; Length 2910;
Best Local Similarity 77.8%; Pred. No. 2.9e+03;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LORPGLO 9
DB 2853 LLMRGLRL 2861

RESULT 12
US-09-739-449-11020
Sequence 11020, Application US/09739449
GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15490)C
CURRENT APPLICATION NUMBER: US/09/739,449
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 11020
LENGTH: 291
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11020

Query Match 53.3%; Score 32.5; DB 5; Length 291;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 4; Mismatches 0; Indels 3; Gaps 1;

QY 2 LORPGLO--LYFS 12
DB 54 VERPGLOVARVFFS 67

RESULT 13
US-09-724-475-3
Sequence 3, Application US/09724475
GENERAL INFORMATION:
APPLICANT: Tsarev, Sergei A., Emerson,
Suzanne U., Purcell, Robert H.
TITLE OF INVENTION: Recombinant Proteins Of
A Pakistani Strain Of Hepatitis E And Their
Use In Diagnostic Methods And Vaccines
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/724,475
FILING DATE: 28-Nov-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,523
FILING DATE: <Unknown>
APPLICATION NUMBER: US08/316,765
FILING DATE: 03-OCT-1994
APPLICATION NUMBER: 07/947,263
FILING DATE: 18-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4032054
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acid residues
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-724-475-3

Query Match 52.5%; Score 32; DB 5; Length 123;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RGLGLYFS 12
DB 83 RGLDLVFA 91

RESULT 14
US-09-769-066-21
Sequence 21, Application US/09769066
GENERAL INFORMATION:
APPLICANT: Fuerst, Thomas R.
Mcatee, C. Patrick
Yarough, Patrice O.
Zhang, Yifan
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 21:


```

SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: Hepatitis E virus (Burma Strain)
ORF-3
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-769-066-21
    
```

```

Query Match          52.5%; Score 32; DB 5; Length 124;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 RPLQLYFS 12
   |||||
Db 83 RPLGLVFA 91
    
```

```

RESULT 15
US-09-270-849B-191061
Sequence 191061, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 191061
LENGTH: 141
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-191061
    
```

```

Query Match          52.5%; Score 32; DB 5; Length 141;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 LQRPGLQ 8
   |||||
Db 66 LQRPGLQ 73
    
```

Search completed: June 4, 2001, 12:21:29
 Job time: 591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: June 4, 2001, 12:20:02 ; Search time 260.2 seconds
(without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-24
Perfect score: 64
Sequence: 1 FYMSKTLRGN 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5 -

Searched: 1009251 seqs, 160854530 residues
Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents: AA Main: *
1: /cgn2_6/ptodata/2/paa/pctus.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US60.COMB.pep.*

Pref. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	64	100.0	58	US-60-160-209-3193
2	64	100.0	96	US-60-160-209-4252
3	64	100.0	102	US-60-160-203-4414
4	64	100.0	261	US-09-245-764-7
5	64	100.0	374	US-09-702-021-14
6	64	100.0	399	PCT-US00-05882-1128
7	40	62.5	91	US-08-332-562-94
8	40	62.5	91	US-08-462-973-94
9	40	62.5	410	US-07-861-895-25
10	38	59.4	268	US-09-634-238-323

11	38	59.4	720	US-09-634-238-322	Sequence 322, App
12	37	57.8	197	US-09-417-507-39462	Sequence 39462, A
13	37	57.8	1845	US-60-245-201-234	Sequence 234, App
14	36	56.2	143	US-09-248-796-26957	Sequence 26957, A
15	36	56.2	153	PCT-US01-01324-2181	Sequence 2181, App
16	36	56.2	356	US-60-167-217-23150	Sequence 23150, A
17	36	56.2	419	US-09-270-767-45469	Sequence 45469, A
18	36	56.2	514	US-60-167-217-23147	Sequence 23147, A
19	36	56.2	669	US-60-167-217-23115	Sequence 23115, A
20	36	56.2	669	US-60-167-217-23174	Sequence 23174, A
21	36	56.2	669	US-60-173-464-18719	Sequence 18719, A
22	36	56.2	669	US-60-173-464-18758	Sequence 18758, A
23	36	56.2	669	US-60-191-637-22838	Sequence 22838, A
24	36	56.2	669	US-60-191-637-22883	Sequence 22883, A
25	36	56.2	669	US-60-191-681-18005	Sequence 18005, A
26	36	56.2	669	US-60-191-681-18038	Sequence 18038, A
27	36	56.2	671	US-60-173-464-18737	Sequence 18737, A
28	36	56.2	671	US-60-191-637-22862	Sequence 22862, A
29	36	56.2	671	US-60-191-637-22820	Sequence 18020, A
30	36	56.2	791	US-60-155-811-561	Sequence 561, App
31	36	56.2	791	US-60-155-811-801	Sequence 801, App
32	36	56.2	791	US-60-155-811-893	Sequence 893, App
33	36	56.2	1015	US-09-417-507-39489	Sequence 39489, A
34	35	54.7	84	US-09-417-507-24640	Sequence 24640, A
35	35	54.7	96	US-60-196-190-1076	Sequence 1076, App
36	35	54.7	101	US-60-164-762-674	Sequence 674, App
37	35	54.7	107	US-60-170-430-2622	Sequence 2622, App
38	35	54.7	121	US-09-417-507-36567	Sequence 36567, A
39	35	54.7	126	US-60-170-430-2462	Sequence 2462, App
40	35	54.7	137	US-09-417-507-42536	Sequence 42536, A
41	35	54.7	147	US-60-177-646-2322	Sequence 2322, App
42	35	54.7	210	US-60-188-162-3505	Sequence 3505, App
43	35	54.7	237	US-60-185-361-508	Sequence 508, App
44	35	54.7	269	US-60-207-216-747	Sequence 747, App
45	35	54.7	309	US-60-185-362-722	Sequence 722, App

ALIGNMENTS

RESULT 1
US-60-160-209-3193
Sequence 3193, Application US/60160209
GENERAL INFORMATION:
APPLICANT: BONAZZI, VITEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL000113
CURRENT APPLICATION NUMBER: US/60/160,209
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3193
LENGTH: 58
TYPE: PRT
ORGANISM: HUMAN
US-60-160-209-3193

Query Match 100.0%; Score 64; DB:23; Length 58;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches: 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYMSKTLRGN 12
|||||
Db 25 FYMSKTLRGN 36

RESULT 2
US-60-160-209-4252
Sequence 4252, Application US/60160209
GENERAL INFORMATION:

APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
FILE REFERENCE: CL000113
CURRENT APPLICATION NUMBER: US/60/160,209
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4252
LENGTH: 96
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(96)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-209-4252

Query Match
Best Local Similarity: 100.0%; Score 64; DB 23; Length 96;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
DB 43 FYMGSKTLGRN 54

RESULT 3
US-60-160-203-4414
Sequence 4414, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: CL000116
CURRENT APPLICATION NUMBER: US/60/160,203
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4414
LENGTH: 102
TYPE: PRT
ORGANISM: HUMAN
US-60-160-203-4414

Query Match
Best Local Similarity: 100.0%; Score 64; DB 23; Length 102;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
DB 51 FYMGSKTLGRN 62

RESULT 4
US-09-245-764-7
Sequence 7, Application US/09245764
GENERAL INFORMATION:
APPLICANT: HOGARTH, P. MARK
APPLICANT: POWELL, MARIE S.
APPLICANT: MCKENZIE, IAN F.C.
APPLICANT: MAXWELL, KELLY F.
APPLICANT: GARRETT, THOMAS P.J.
APPLICANT: EPA, VIDANA
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES AND MODELS OF FC RECEPTORS
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 4102-4
CURRENT APPLICATION NUMBER: US/09/245,764
CURRENT FILING DATE: 1999-02-05

EARLIER APPLICATION NUMBER: 60/099,994
EARLIER FILING DATE: 1998-09-11
EARLIER APPLICATION NUMBER: 60/073,972
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 261
TYPE: PRT
ORGANISM: Homo sapiens
US-09-245-764-7

Query Match
Best Local Similarity: 100.0%; Score 64; DB 16; Length 261;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
DB 210 FYMGSKTLGRN 221

RESULT 5
US-09-702-021-14
Sequence 14, Application US/09702021
GENERAL INFORMATION:
APPLICANT: C. FRASER
TITLE OF INVENTION: PAIR MOLECULES AND USES THEREOF
FILE REFERENCE: 7853-217
CURRENT APPLICATION NUMBER: US/09/702,021
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 14
LENGTH: 374
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-021-14

Query Match
Best Local Similarity: 100.0%; Score 64; DB 21; Length 374;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
DB 229 FYMGSKTLGRN 240

RESULT 6
PCT-US00-05882-1128
Sequence 1128, Application PC/TUS0005882
GENERAL INFORMATION:
APPLICANT: CRAIG ROSEN,
APPLICANT: STEVE RUBEN
TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
FILE REFERENCE: PA106PCT
CURRENT APPLICATION NUMBER: PCT/US00/05882
CURRENT FILING DATE: 2000-03-08
EARLIER APPLICATION NUMBER: 60/124,270
EARLIER FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1128
LENGTH: 399
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (208)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE

LOCATION: (349)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1128

Query Match 100.0%; Score 64; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYGSKTLGRN 12
|||||
DB 254 FYGSKTLGRN 265

RESULT 7
US-08-332-562-94
Sequence 94, Application US/08332562
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562-94

Query Match 62.5%; Score 40; DB 7; Length 91;
Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYGSKTLGRN 12
|||||
DB 41 FYGSKTLGRN 52

RESULT 8
US-08-462-973-94

Sequence 94, Application US/08462973
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,973
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/332,562
FILING DATE: 31-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM8232/94
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU87/00159
FILING DATE: 29-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH6166
FILING DATE: 26-MAY-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-462-973-94

Query Match 62.5%; Score 40; DB 8; Length 91;
Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYGSKTLGRN 12
|||||
DB 41 FYGSKTLGRN 52

RESULT 9
US-07-861-895-25
Sequence 25, Application US/07861895
GENERAL INFORMATION:
APPLICANT: HOGARTH, Phillip M.
APPLICANT: HULETT, Mark D.

APPLICANT: IERINO, Francesco L.
APPLICANT: MCKENZIE, Ian F. C.
APPLICANT: OSMAN, Nalin
TITLE OF INVENTION: HYBRID FC RECEPTOR MOLECULES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/861,895
FILING DATE: 19920624
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00513
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7045
FILING DATE: 25-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17227/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-861-895-25

Query Match
Best Local Similarity 62.5%; Score 40; DB 3; Length 410;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYGSKTLGRN 12
DB 244 FYGSKTLEARN 255

RESULT 10
US-09-634-238-323
Sequence 323, Application US/09634238
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
APPLICANT: Dekker, James
APPLICANT: Christensson, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
TITLE OF INVENTION: Polynucleotides, materials incorporating
TITLE OF INVENTION: them and methods for using them
FILE REFERENCE: 11000.104301
CURRENT APPLICATION NUMBER: US/09/634,238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 323
LENGTH: 268
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-09-634-238-323

Query Match
Best Local Similarity 59.4%; Score 38; DB 20; Length 268;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YMGSKTLGRN 12
DB 243 YVGOKPLRSRN 253

RESULT 11
US-09-634-238-322
Sequence 322, Application US/09634238
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
APPLICANT: Dekker, James
APPLICANT: Christensson, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
TITLE OF INVENTION: Polynucleotides, materials incorporating
TITLE OF INVENTION: them and methods for using them
FILE REFERENCE: 11000.104301
CURRENT APPLICATION NUMBER: US/09/634,238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 322
LENGTH: 720
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-09-634-238-322

Query Match
Best Local Similarity 59.4%; Score 38; DB 20; Length 720;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YMGSKTLGRN 12
DB 407 YVGOKPLRSRN 417

RESULT 12
US-09-417-507-39462
Sequence 39462, Application US/09417507
GENERAL INFORMATION:
APPLICANT: KEITH G. WEINSTOCK ET AL.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
FILE REFERENCE: PATH99-10
CURRENT APPLICATION NUMBER: US/09/417,507
CURRENT FILING DATE: 1999-10-14
NUMBER OF SEQ ID NOS: 44312
SEQ ID NO 39462
LENGTH: 197
TYPE: PRT
ORGANISM: A. fumigatus
US-09-417-507-39462

Query Match
Best Local Similarity 57.8%; Score 37; DB 18; Length 197;
Matches 45.5%; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGSKTLGRN 12

Db 75 FLGGRTIRGRS 85

RESULT 13

US-60-245-201-234

Sequence 234, Application US/60245201

GENERAL INFORMATION:

APPLICANT: Beasley, Ellen

TITLE OF INVENTION: ISOLATED HUMAN PHASE II

TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING

FILE REFERENCE: CL000879

CURRENT APPLICATION NUMBER: US/60/245,201

CURRENT FILING DATE: 2000-11-03

NUMBER OF SEQ ID NOS: 381

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 234

LENGTH: 1845

TYPE: PRT

ORGANISM: Human

US-60-245-201-234

Query Match 57.8%; Score 37; DB 23; Length 1845;

Best Local Similarity 66.7%; Pred. No. 5.6e+02;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12

Db 835 FTPEGQTERGRN 846

RESULT 14

US-09-248-796-26957

Sequence 26957, Application US/09248796

GENERAL INFORMATION:

APPLICANT: Keith Weinstein et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796

CURRENT FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 28206

SEQ ID NO 26957

LENGTH: 143

TYPE: PRT

ORGANISM: Candida albicans

US-09-248-796-26957

Query Match 56.2%; Score 36; DB 16; Length 143;

Best Local Similarity 63.8%; Pred. No. 52;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYMGSKTLGR 11

Db 89 FKIGSKTLEGO 99

RESULT 15

PCT-US01-01324-2181

Sequence 2181, Application PC/TUS0101324

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc., et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC002PCT

CURRENT APPLICATION NUMBER: PCT/US01/01324

CURRENT FILING DATE: 2001-01-14

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 5116

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2181

LENGTH: 153

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (34)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (55)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (64)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (68)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (70)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

PCT-US01-01324-2181

Query Match 56.2%; Score 36; DB 1; Length 153;

Best Local Similarity 87.5%; Pred. No. 56;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GSKTLGR 11

Db 98 GSQTLGR 105

Search completed: June 4, 2001, 12:20:03
Job time: 516 sec

Tue Jun 5 07:08:55 2001

us-09-284-107-24.rapm

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:21:29 ; Search time 84.28 Seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-24
Perfect score: 64
Sequence: 1 FYMSGKTLGRN 12

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits, satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Pending_Patents_AA.New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	12	US-09-284-107-24	Sequence 24, App1
2	37	57.8	212	US-09-459-573-12	Sequence 12, App1
3	35	54.7	61	PCT-US00-01979A-69	Sequence 69, App1
4	35	54.7	61	PCT-US00-01979A-73	Sequence 73, App1
5	35	54.7	61	US-09-493-795A-69	Sequence 69, App1
6	35	54.7	61	US-09-493-795A-73	Sequence 73, App1
7	35	54.7	64	PCT-US00-01979A-65	Sequence 65, App1
8	35	54.7	64	US-09-493-795A-65	Sequence 65, App1
9	35	54.7	377	US-09-739-449-12089	Sequence 12089, A
10	34	53.1	334	US-09-739-449-12004	Sequence 12004, A
11	33	51.6	55	US-09-270-849B-194148	Sequence 194148, A
12	33	51.6	275	US-09-823-356-16	Sequence 16, App1
13	33	51.6	282	PCT-US01-11988-458	Sequence 458, App
14	33	51.6	282	US-09-833-245-458	Sequence 458, App
15	33	51.6	291	US-09-525-361-29	Sequence 29, App1
16	33	51.6	318	PCT-US01-11988-455	Sequence 455, App
17	33	51.6	318	US-09-833-245-455	Sequence 455, App
18	33	51.6	553	US-09-739-449-11123	Sequence 11123, A
19	32.5	50.8	546	US-09-603-124B-410	Sequence 410, App
20	32	50.0	61	PCT-US00-01979A-71	Sequence 71, App1
21	32	50.0	61	US-09-493-795A-71	Sequence 71, App1
22	32	50.0	140	US-09-612-795-1	Sequence 1, App1
23	32	50.0	140	US-09-612-795-2	Sequence 2, App1
24	32	50.0	140	US-09-612-795-3	Sequence 3, App1
25	32	50.0	140	US-09-612-795-4	Sequence 4, App1
26	32	50.0	411	US-09-403-736-2	Sequence 2, App1
27	32	50.0	670	PCT-US01-04098A-1054	Sequence 1054, App

28	32	50.0	749	1	PCT-US01-04098A-3022	Sequence 3022, Ap
29	31	48.4	110	5	US-09-270-849B-193304	Sequence 193304, A
30	31	48.4	177	5	US-09-270-849B-194847	Sequence 194847, A
31	31	48.4	238	5	US-09-270-849B-167832	Sequence 167832, A
32	31	48.4	241	5	US-09-739-449-9493	Sequence 9493, Ap
33	31	48.4	260	5	US-09-696-382-139	Sequence 139, App
34	31	48.4	387	5	US-09-739-449-11227	Sequence 11227, A
35	31	48.4	475	5	US-09-739-449-11741	Sequence 11741, A
36	31	48.4	681	1	PCT-US01-01900-2	Sequence 2, App1
37	31	48.4	842	5	US-09-739-449-10658	Sequence 10658, A
38	31	48.4	959	5	US-09-270-849B-190396	Sequence 190396, A
39	31	48.4	1190	1	PCT-US01-04098A-1015	Sequence 1015, Ap
40	31	48.4	1210	1	PCT-US01-04098A-2983	Sequence 2983, Ap
41	30	46.9	64	1	PCT-US01-11988-1005	Sequence 1005, Ap
42	30	46.9	64	5	US-09-833-245-1005	Sequence 1005, Ap
43	30	46.9	93	5	US-09-270-849B-189576	Sequence 189576, A
44	30	46.9	201	5	US-09-826-019-63	Sequence 63, App1
45	30	46.9	204	5	US-09-696-382-147	Sequence 147, App

ALIGNMENTS

```
RESULT 1
US-09-284-107-24
Sequence 24, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
FILE REFERENCE: 31363200600
CURRENT APPLICATION NUMBER: US/09/284,107
PRIOR FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
US-09-284-107-24
Query Match 100.0%; Score 64; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 FYMSGKTLGRN 12
RESULT 2
US-09-459-573-12
Sequence 12, Application US/09459573
GENERAL INFORMATION:
APPLICANT: LIVSHITS, VITALITY
APPLICANT: ZAKATAEVA, NATALIA
APPLICANT: NAKANISHI, KAZUO
APPLICANT: VENIAMINOVICH, VLADIMIR
APPLICANT: TROSHIN, PETR
APPLICANT: TOKHAKOVA, IRINA
TITLE OF INVENTION: METHOD FOR PRODUCING L-AMINO ACIDS
FILE REFERENCE: 0010-1066-0
CURRENT APPLICATION NUMBER: US/09/459,573
CURRENT FILING DATE: 1999-12-13
```


RESULT 7
PCT-US00-01979A-65
; Sequence 65, Application PC/TUS0001979A
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; TITLE OF INVENTION: Alpha-Conotoxin Peptides
; FILE REFERENCE: Alphas 2
; CURRENT APPLICATION NUMBER: PCT/US00/01979A
; CURRENT FILING DATE: 2000-01-28
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 404
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 65
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Conus textile
PCT-US00-01979A-65

Query Match 54.7%; Score 35; DB 1; Length 64;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12
| | | | |
DB 17 FSSGRSTFRGRN 28

RESULT 8
US-09-493-795A-65
; Sequence 65, Application US/09493795A
; GENERAL INFORMATION:
; APPLICANT: Watkins, Maren
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Hillyard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Alpha-Conotoxin Peptides
; FILE REFERENCE: 2314-179 A
; CURRENT APPLICATION NUMBER: US/09/493,795A
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/118,381
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 404
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 65
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Conus textile
US-09-493-795A-65

Query Match 54.7%; Score 35; DB 5; Length 64;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12
| | | | |
DB 17 FSSGRSTFRGRN 28

RESULT 9
US-09-739-449-12089
; Sequence 12089, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C

CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 12089
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-12089

Query Match 54.7%; Score 35; DB 5; Length 377;
Best Local Similarity 58.3%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12
| | | | |
DB 302 FYLGSAITRSRRN 313

RESULT 10
US-09-739-449-12004
; Sequence 12004, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 12004
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-12004

Query Match 53.1%; Score 34; DB 5; Length 334;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 YMGSKTLGR 11
| | | | |
DB 296 HMGSAITRGR 305

RESULT 11
US-09-270-849B-194148
; Sequence 194148, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 194148
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: at any location having Xaa, Yaa means any amino acid
US-09-270-849B-194148

Query Match 51.6%; Score 33; DB 5; Length 55;
Best Local Similarity 54.5%; Pred. No. 24;

RESULT 12
US-09-823-356-16
Sequence 16, Application US/09823356
GENERAL INFORMATION:
APPLICATION: 09823356

```

APPLICANT: Yang, Y. Tom
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Kaser, Matthew R.
APPLICANT: Baughn, Mariah R.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN MEMBRANE SPANNING PROTEINS
FILE REFERENCE: PR-0489-1 CON
CURRENT APPLICATION NUMBER: US/09/823,356
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/039,307
PRIOR FILING DATE: 1998 March 13
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PERL Program
SEQ ID NO 16
LENGTH: 275
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID NO: 2879792
US-09-823-356-16

```

Query Match	51.6%;	Score 33;	DB 5;	Length 275;
Best Local Similarity	85.7%;	Pred. NO. 1.3e+02;		
Matches	6;	Conservative	1;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	2	YMGSKTL	8
		1:11111	
Db	39	YLGSKTL	45

```

RESULT 13
PCT-US01-11988-458
Sequence 458, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCP
CURRENT APPLICATION NUMBER: PCT/US01/11988
PRIORITY FILING DATE: 2001-01-12
PRIORITY APPLICATION NUMBER: 60/229,358
PRIORITY FILING DATE: 2000-04-12
PRIORITY APPLICATION NUMBER: 60/256,931
PRIORITY FILING DATE: 2000-12-21
PRIORITY APPLICATION NUMBER: 60/199,384
PRIORITY FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 458
LENGTH: 282
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-458

```

Query Match	51.6%;	Score 33;	DB 1;	Length 282;
-------------	--------	-----------	-------	-------------

RESULT 14
US-09-833-245-458

? Sequence 458, Application US/09833245
 ? GENERAL INFORMATION:
 ? APPLICANT: Human Genome Sciences, Inc.
 ? TITLE OF INVENTION: Albumin Fusion Proteins
 ? FILE REFERENCE: PFS46PCP
 ? CURRENT APPLICATION NUMBER: US/09/833,245
 ? CURRENT FILING DATE: 2001-04-12
 ? PRIOR APPLICATION NUMBER: 60/229, 358
 ? PRIOR FILING DATE: 2000-04-12
 ? PRIOR APPLICATION NUMBER: 60/256, 931
 ? PRIOR FILING DATE: 2000-12-21
 ? PRIOR APPLICATION NUMBER: 60/199, 384
 ? PRIOR FILING DATE: 2000-04-25
 ? NUMBER OF SEQ ID NOS: 2267
 ? SOFTWARE: PatentIn Ver. 2.1
 ? SEQ ID NO 458
 ? LENGTH: 282
 ? TYPE: ERT
 ? ORGANISM: Homo sapiens
 ? OS-09-833-245-458

Query Match	51.6%;	Score 33;	DB 5;	Length 282;
Best Local Similarity	85.7%;	Pred. No. 1.3e+02;		
Matches	6;	Conservative	1;	Mismatches 0;
			Indels	0;
			Gaps	0

QY	2	YMGSKTL	8
		:	
Db	46	YLGSKTL	52

```

RESULT 15
US-09-525-361-29
Sequence 29, Application US/09525361
GENERAL INFORMATION:
APPLICANT: MACK, DAVID
TITLE OF INVENTION: KURT
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING AND TREATING BREAST CANCER
TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF SCREENING FOR BREAST
FILE REFERENCE: A-67860-3/DJB/JJD
CURRENT APPLICATION NUMBER: US/09/525,361
PRIORITY FILING DATE: 2000-03-15
PRIORITY APPLICATION NUMBER: US 09/268,865
PRIORITY FILING DATE: 1999-03-15
PRIORITY APPLICATION NUMBER: US 09/450,810
PRIORITY FILING DATE: 1999-11-29
PRIORITY APPLICATION NUMBER: US 09/453,137
PRIORITY FILING DATE: 1999-12-02
PRIORITY APPLICATION NUMBER: US 09/439,878
PRIORITY FILING DATE: 1999-11-12
PRIORITY APPLICATION NUMBER: US 09/440,370
PRIORITY FILING DATE: 1999-11-12
PRIORITY APPLICATION NUMBER: US 09/440,493
PRIORITY FILING DATE: 1999-11-15
PRIORITY APPLICATION NUMBER: US 09/520,478
PRIORITY FILING DATE: 2000-03-08
PRIORITY APPLICATION NUMBER: US 09/440,676
PRIORITY FILING DATE: 1999-11-16
PRIORITY APPLICATION NUMBER: US 09/440,677
PRIORITY FILING DATE: 1999-11-16
NUMBER OF SEQ ID NOS: 53
SOFTWARE: Patentlin Ver. 2.1

```

SEQ ID NO 29
 LENGTH: 291
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-525-361-29

Query Match 51.6%; Score 33; DB 5; Length 291;
 Best Local Similarity 85.7%; Pred. NO. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YMSKTL 8
 1:|||||
 Db 55 YLSKTL 61

Search completed: June 4, 2001, 12:21:29
 Job time: 591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:20:03 ; Search time 260.2 seconds
(without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-25
Perfect score: 57
Sequence: 1 TSSEYQILFARR 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues
Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/2/paa/US06.COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US09.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	57	100.0	58	US-60-160-209-3193
2	57	100.0	96	US-60-160-209-4252
3	57	100.0	102	US-60-160-209-4414
4	57	100.0	261	US-09-245-764-7
5	57	100.0	374	US-09-702-021-14
6	57	100.0	399	PCT-US00-05882-1128
7	36	63.2	56	US-60-173-468-1209
8	36	63.2	23	US-60-181-426-621
9	36	63.2	91	US-08-332-562-94
10	36	63.2	91	US-08-462-973-94

11	36	63.2	151	23	US-60-207-216-686	Sequence 686, App
12	36 <td>63.2</td> <td>325</td> <td>18</td> <td>US-09-489-039A-8221</td> <td>Sequence 8221, App</td>	63.2	325	18	US-09-489-039A-8221	Sequence 8221, App
13	36 <td>63.2</td> <td>366</td> <td>21</td> <td>US-09-718-692-4</td> <td>Sequence 4, Appl</td>	63.2	366	21	US-09-718-692-4	Sequence 4, Appl
14	36 <td>63.2</td> <td>366</td> <td>21</td> <td>US-09-718-815-4</td> <td>Sequence 4, Appl</td>	63.2	366	21	US-09-718-815-4	Sequence 4, Appl
15	36 <td>63.2</td> <td>366</td> <td>21</td> <td>US-09-718-852-4</td> <td>Sequence 4, Appl</td>	63.2	366	21	US-09-718-852-4	Sequence 4, Appl
16	36 <td>63.2</td> <td>410</td> <td>3</td> <td>US-07-861-895-25</td> <td>Sequence 25, Appl</td>	63.2	410	3	US-07-861-895-25	Sequence 25, Appl
17	36 <td>63.2</td> <td>417</td> <td>16</td> <td>US-09-252-999A-16876</td> <td>Sequence 16876, A</td>	63.2	417	16	US-09-252-999A-16876	Sequence 16876, A
18	36 <td>63.2</td> <td>1637</td> <td>21</td> <td>US-09-718-692-2</td> <td>Sequence 2, Appl</td>	63.2	1637	21	US-09-718-692-2	Sequence 2, Appl
19	36 <td>63.2</td> <td>1637</td> <td>21</td> <td>US-09-718-815-2</td> <td>Sequence 2, Appl</td>	63.2	1637	21	US-09-718-815-2	Sequence 2, Appl
20	36 <td>63.2</td> <td>1637</td> <td>21</td> <td>US-09-718-852-2</td> <td>Sequence 2, Appl</td>	63.2	1637	21	US-09-718-852-2	Sequence 2, Appl
21	36 <td>63.2</td> <td>2004</td> <td>16</td> <td>US-09-231-899-9</td> <td>Sequence 9, Appl</td>	63.2	2004	16	US-09-231-899-9	Sequence 9, Appl
22	35 <td>61.4</td> <td>104</td> <td>23</td> <td>US-60-173-464-25166</td> <td>Sequence 25166, A</td>	61.4	104	23	US-60-173-464-25166	Sequence 25166, A
23	35 <td>61.4</td> <td>104</td> <td>23</td> <td>US-60-191-637-32284</td> <td>Sequence 32284, A</td>	61.4	104	23	US-60-191-637-32284	Sequence 32284, A
24	35 <td>61.4</td> <td>104</td> <td>23</td> <td>US-60-191-681-25591</td> <td>Sequence 25591, A</td>	61.4	104	23	US-60-191-681-25591	Sequence 25591, A
25	35 <td>61.4</td> <td>112</td> <td>18</td> <td>US-09-417-507-36446</td> <td>Sequence 36446, A</td>	61.4	112	18	US-09-417-507-36446	Sequence 36446, A
26	35 <td>61.4</td> <td>553</td> <td>23</td> <td>US-60-167-217-7951</td> <td>Sequence 7951, App</td>	61.4	553	23	US-60-167-217-7951	Sequence 7951, App
27	35 <td>61.4</td> <td>553</td> <td>23</td> <td>US-60-173-464-6355</td> <td>Sequence 6355, App</td>	61.4	553	23	US-60-173-464-6355	Sequence 6355, App
28	35 <td>61.4</td> <td>553</td> <td>23</td> <td>US-60-191-637-7854</td> <td>Sequence 7854, App</td>	61.4	553	23	US-60-191-637-7854	Sequence 7854, App
29	35 <td>61.4</td> <td>553</td> <td>23</td> <td>US-60-191-681-6123</td> <td>Sequence 6123, App</td>	61.4	553	23	US-60-191-681-6123	Sequence 6123, App
30	34	59.6	174	15	US-09-107-533-4715	Sequence 4715, App
31	34	59.6	220	18	US-09-417-507-37264	Sequence 37264, A
32	34	59.6	301	18	US-09-489-039A-12472	Sequence 12472, A
33	34	59.6	341	16	US-09-248-796-14902	Sequence 14902, A
34	33	57.9	94	23	US-60-162-248-932	Sequence 932, App
35	33	57.9	94	23	US-60-169-867-6400	Sequence 6400, App
36	33	57.9	193	1	PCT-US97-05223-678	Sequence 678, App
37	33	57.9	193	11	US-08-761-318-678	Sequence 678, App
38	33	57.9	193	12	US-08-824-132-678	Sequence 678, App
39	33	57.9	193	13	US-08-993-002A-5167	Sequence 5167, App
40	33	57.9	223	23	US-60-185-362-615	Sequence 615, App
41	33	57.9	225	1	PCT-US00-15135-152	Sequence 152, App
42	33	57.9	229	1	PCT-US00-15135-151	Sequence 151, App
43	33	57.9	239	17	US-09-328-352-5274	Sequence 5274, App
44	33	57.9	293	16	US-09-252-991A-21673	Sequence 21673, A
45	33	57.9	295	1	PCT-US00-15135-149	Sequence 149, App

ALIGNMENTS

RESULT 1
US-60-160-209-3193
; Sequence 3193, Application US/60160209
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID-MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CLO00113
; CURRENT APPLICATION NUMBER: US/60/160, 209
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4646
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3193
; LENGTH: 58
; TYPE: PRT
; ORGANISM: HUMAN
; US-60-160-209-3193

Query Match 100.0%; Score 57; DB 23; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSSEYQILFARR 12
| | | | | | | | | | | | | |
Db 37 TSSEYQILFARR 48

RESULT 2
US-60-160-209-4252
; Sequence 4252, Application US/60160209
; GENERAL INFORMATION:

APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: C1000113
CURRENT APPLICATION NUMBER: US/60/160,209
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4252
LENGTH: 96
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)-(96)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-209-4252

Query Match 100.0%; Score 57; DB 23; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
|||||
DB 55 TSSEYOILTARR 66

RESULT 3
US-60-160-203-4414
Sequence 4414, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: C1000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4414
LENGTH: 102
TYPE: PRT
ORGANISM: HUMAN
US-60-160-203-4414

Query Match 100.0%; Score 57; DB 23; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
|||||
DB 63 TSSEYOILTARR 74

RESULT 4
US-09-245-764-7
Sequence 7, Application US/09245764
GENERAL INFORMATION:
APPLICANT: HOGARTH, P. MARK
APPLICANT: HOGARTH, P. MARK
APPLICANT: POWELL, MARIE S.
APPLICANT: MCKENZIE, IAN F.C.
APPLICANT: GARRETT, THOMAS P.J.
APPLICANT: EPPA, VIDANA
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES AND MODELS OF PC RECEPTORS
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 4102-4
CURRENT APPLICATION NUMBER: US/09/245,764
CURRENT FILING DATE: 1999-02-05

EARLIER APPLICATION NUMBER: 60/099,994
EARLIER FILING DATE: 1998-09-11
EARLIER APPLICATION NUMBER: 60/073,972
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 261
TYPE: PRT
ORGANISM: Homo sapiens
US-09-245-764-7

Query Match 100.0%; Score 57; DB 16; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
|||||
DB 222 TSSEYOILTARR 233

RESULT 5
US-09-702-021-14
Sequence 14, Application US/09702021
GENERAL INFORMATION:
APPLICANT: C. FRASER
TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF
FILE REFERENCE: 7853-217
CURRENT APPLICATION NUMBER: US/09/702,021
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 14
LENGTH: 374
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-021-14

Query Match 100.0%; Score 57; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
|||||
DB 241 TSSEYOILTARR 252

RESULT 6
PCT-US00-05882-1128
Sequence 1128, Application PC/TUS0005882
GENERAL INFORMATION:
APPLICANT: CRAIG ROSEN,
APPLICANT: STEVE RUBEN
TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
FILE REFERENCE: PA106PCM
CURRENT APPLICATION NUMBER: PCT/US00/05882
CURRENT FILING DATE: 2000-03-08
EARLIER APPLICATION NUMBER: 60/124,270
EARLIER FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1128
LENGTH: 399
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (208)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE

LOCATION: (349)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1128

Query Match
Best Local Similarity 100.0%; Score 57; DB 1; Length 399;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYQILTRR 12
Db 266 TSSEYQILTRR 277

RESULT 7
US-60-173-468-1209
Sequence 1209, Application US/60173468
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
FILE REFERENCE: CL000181
CURRENT APPLICATION NUMBER: US/60/173,468
CURRENT FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 1418
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1209
LENGTH: 56
TYPE: PRT
ORGANISM: HUMAN
US-60-173-468-1209

Query Match
Best Local Similarity 70.0%; Score 36; DB 23; Length 56;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 SSEYQILTRR 11
Db 15 SSEYQILTRR 24

RESULT 8
US-60-181-426-621
Sequence 621, Application US/60181426
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
FILE REFERENCE: CL000212
CURRENT APPLICATION NUMBER: US/60/181,426
CURRENT FILING DATE: 2000-02-09
NUMBER OF SEQ ID NOS: 708
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 621
LENGTH: 56
TYPE: PRT
ORGANISM: HUMAN
US-60-181-426-621

Query Match
Best Local Similarity 63.2%; Score 36; DB 23; Length 56;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 SSEYQILTRR 11
Db 15 SSEYQILTRR 24

RESULT 9

US-08-332-562-94
Sequence 94, Application US/08332562

GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-332-562-94

Query Match
Best Local Similarity 66.7%; Score 36; DB 7; Length 91;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TSSEYQILTRR 12
Db 53 TSSEYQILTRR 64

RESULT 10
US-08-462-973-94
Sequence 94, Application US/08462973
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,973
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/332,562
FILING DATE: 31-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM8232/94
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU87/00159
FILING DATE: 29-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH6166
FILING DATE: 26-MAY-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-462-973-94

Query Match 63.2%; Score 36; DB 8; Length 91;
Best Local Similarity 66.7%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TSSEYOILTARR 12
DB 53 TSSEYHARRER 64

RESULT 11
US-60-207-216-686
Sequence 686, Application US/60207216
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
FILE REFERENCE: CLO00580
CURRENT APPLICATION NUMBER: US/60/207,216
CURRENT FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 898
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 686
LENGTH: 151
TYPE: PRT
ORGANISM: HUMAN
US-60-207-216-686

Query Match 63.2%; Score 36; DB 23; Length 151;

Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSEYOILTAR 11
DB 110 SSEYETLTAK 119

RESULT 12
US-09-489-039A-8221
Sequence 8221, Application US/09489039A
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709, 2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8221
LENGTH: 325
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8221

Query Match 63.2%; Score 36; DB 18; Length 325;
Best Local Similarity 58.3%; Pred. No. 42;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSSEYOILTARR 12
DB 182 TLAEYALTLAQR 193

RESULT 13
US-09-718-692-4
Sequence 4, Application US/09718692
GENERAL INFORMATION:
APPLICANT: Berand, Christophe
TITLE OF INVENTION: Novel motor proteins and methods for
FILE REFERENCE: 1052
CURRENT APPLICATION NUMBER: US/09/718,692
CURRENT FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 366
TYPE: PRT
ORGANISM: Human
US-09-718-692-4

Query Match 63.2%; Score 36; DB 21; Length 366;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSEYOILTAR 11
DB 254 SSEYETLTAK 263

RESULT 14
US-09-718-815-4
Sequence 4, Application US/09718815
GENERAL INFORMATION:
APPLICANT: Berand, Christophe
TITLE OF INVENTION: Novel motor proteins and methods for

```

; TITLE OF INVENTION: their use
; FILE REFERENCE: 1052
; CURRENT APPLICATION NUMBER: US/09/718,815
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Human
US-09-718-815-4

```

```

Query Match      63.2%; Score 36; DB 21; Length 366;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY      2 SSEXOILTR 11
Db      254 SSEYETLTR 263

```

```

RESULT 15
US-09-718-852-4
; Sequence 4, Application US/09718852
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1052
; CURRENT APPLICATION NUMBER: US/09/718,852
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Human
US-09-718-852-4

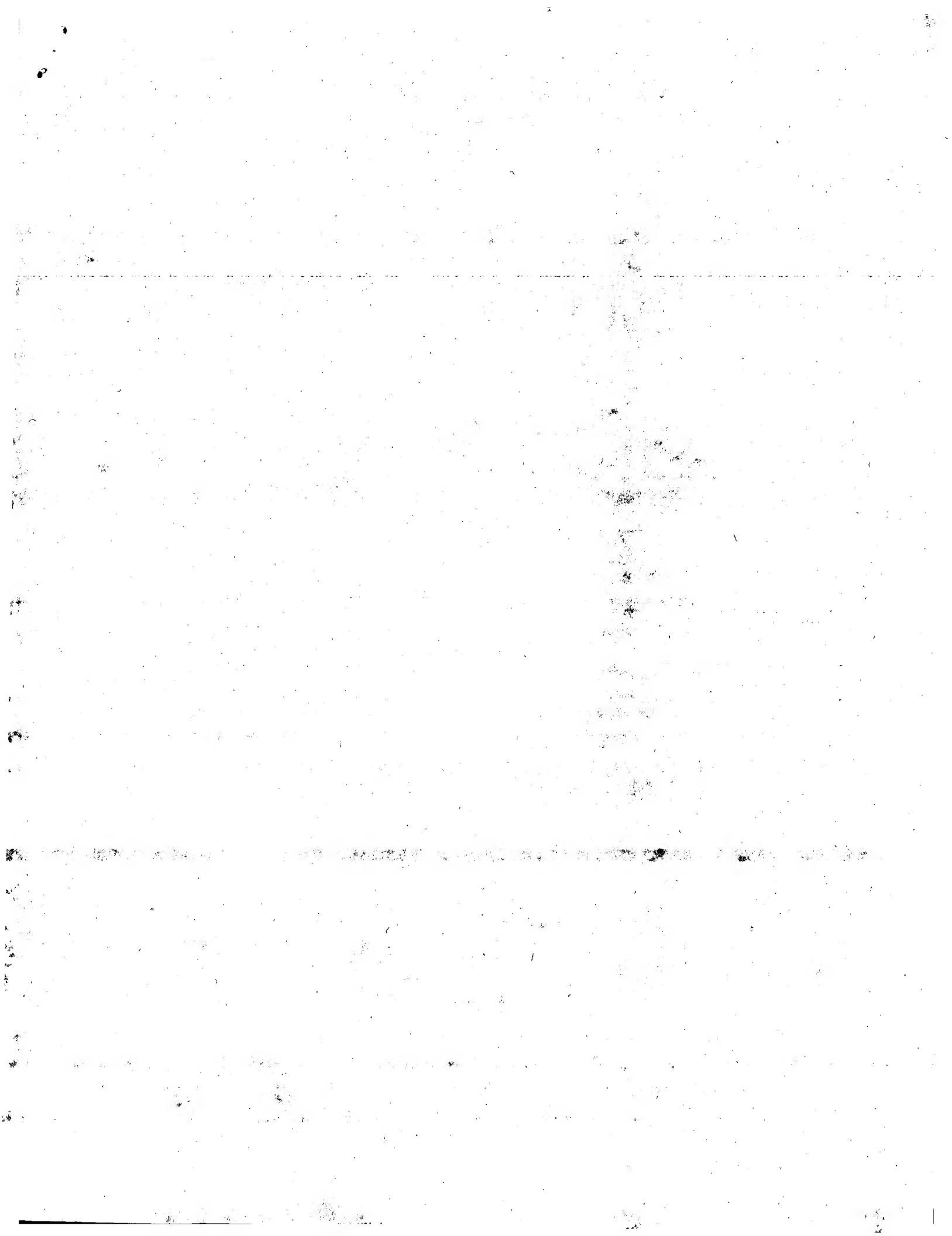
```

```

Query Match      63.2%; Score 36; DB 21; Length 366;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY      2 SSEXOILTR 11
Db      254 SSEYETLTR 263

```

Search completed: June 4, 2001, 12:20:03
 Job time: 516 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:21:29 ; Search time 84.28 Seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-25
Perfect score: 57
Sequence: 1 TSSEXQILTARR 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:
1: /cgn2_6/ptodata1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	12	5	US-09-284-107-25
2	32	56.1	219	1	PCT-US00-35017A-1460
3	32	56.1	236	1	PCT-US01-06769-11
4	32	56.1	245	1	PCT-US01-11988-762
5	32	56.1	245	5	US-09-833-245-762
6	32	56.1	250	1	PCT-US01-11988-763
7	32	56.1	250	5	US-09-833-245-763
8	32	56.1	344	5	US-09-739-449-11547
9	32	56.1	715	6	US-60-248-823-102
10	31	54.4	240	5	US-09-739-449-8297
11	31	54.4	359	1	PCT-US01-11988-133
12	31	54.4	359	5	US-09-833-245-133
13	31	54.4	790	1	PCT-US01-04098A-1378
14	31	54.4	902	5	US-09-193-562D-34
15	31	54.4	4544	1	PCT-US01-04098A-1753
16	31	54.4	4545	5	US-09-750-972-2
17	30	52.6	114	5	US-09-640-211A-721
18	30	52.6	214	5	PCT-US01-04098A-1872
19	30	52.6	240	5	US-09-739-449-12933
20	30	52.6	250	1	PCT-US01-04098A-3840
21	30	52.6	296	5	US-09-270-849B-180859
22	30	52.6	301	5	US-09-270-849B-190000
23	30	52.6	334	5	US-09-739-449-12004
24	30	52.6	344	5	US-09-739-449-8153
25	30	52.6	480	5	US-09-672-459-4
26	30	52.6	483	5	US-09-337-563B-13
27	30	52.6	483	5	US-09-291-023A-16

ALIGNMENTS

28	30	52.6	488	1	PCT-US01-08117-82	Sequence 82, Appl
29	30	52.6	514	5	US-09-337-563B-4	Sequence 4, Appl
30	30	52.6	617	5	US-09-763-702-2	Sequence 2, Appl
31	30	52.6	922	5	US-09-345-236B-89	Sequence 89, Appl
32	30	52.6	1476	6	US-60-248-505-1236	Sequence 1236, Ap
33	29	50.9	144	1	PCT-US01-01312-991	Sequence 991, Ap
34	29	50.9	245	5	US-09-739-449-8568	Sequence 8568, Ap
35	29	50.9	296	5	US-09-270-849B-188275	Sequence 188275, A
36	29	50.9	333	5	US-09-739-449-11488	Sequence 11488, A
37	29	50.9	366	1	PCT-US01-04098A-1697	Sequence 1697, Ap
38	29	50.9	430	1	PCT-US01-04098A-3665	Sequence 3665, Ap
39	29	50.9	438	5	US-09-826-509-581	Sequence 581, Ap
40	29	50.9	495	1	PCT-US01-04098A-3003	Sequence 3003, Ap
41	29	50.9	508	1	PCT-US01-04098A-1035	Sequence 1035, Ap
42	29	50.9	636	5	US-09-739-449-8964	Sequence 8964, Ap
43	29	50.9	854	1	PCT-US01-04098A-1566	Sequence 1566, Ap
44	29	50.9	872	1	PCT-US01-10661-4	Sequence 4, Appl
45	29	50.9	879	5	US-09-459-715-6	Sequence 6, Appl

RESULT 1
US-09-284-107-25
Sequence 25, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: AND PROTEINS HAVING SPECIFIC AFFINITY FOR A TARGET
FILE REFERENCE: 313632000600
CURRENT APPLICATION NUMBER: US/09/284,107
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
PRIOR FILING DATE: 1996-10-08
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
OTHER INFORMATION: domain of CD64
US-09-284-107-25

Query Match 100.0%; Score 57; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 TSSEXQILTARR 12
Db 1 TSSEXQILTARR 12
RESULT 2
PCT-US00-35017A-1460
Sequence 1460, Application PC/TUS0035017A
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 784PC1
CURRENT APPLICATION NUMBER: PCT/US00/35017A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 1478

SEQ ID NO 1460
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-35017A-1460

Query Match 56.1%; Score 32; DB 1; Length 219;
Best Local Similarity 54.5%; Pred. No. 70;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSSEYQILTR 11
|||:|:|:|:
DB 206 TSTDQVLAR 216

RESULT 3
PCT-US01-06769-11
Sequence 11, Application PC/TUS0106769
GENERAL INFORMATION:
APPLICANT: Mayo Medical Ventures
TITLE OF INVENTION: hb7-h2, A NOVEL CO-STIMULATORY MOLECULE
FILE REFERENCE: 07039-202W01
CURRENT APPLICATION NUMBER: PCT/US01/06769
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/186,519
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-06769-11

Query Match 56.1%; Score 32; DB 1; Length 236;
Best Local Similarity 45.5%; Pred. No. 76;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSSEYQILTR 11
|||:|:|:|:
DB 168 TSTDQVLGR 178

RESULT 4
PCT-US01-11988-762
Sequence 762, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patentlin Ver. 2.1
SEQ ID NO 762
LENGTH: 245
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (110)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-11988-762

Query Match 56.1%; Score 32; DB 1; Length 245;
Best Local Similarity 45.5%; Pred. No. 80;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSSEYQILTR 11
|||:|:|:|:
DB 168 TSTDQVLGR 178

RESULT 5
US-09-833-245-762
Sequence 762, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patentlin Ver. 2.1
SEQ ID NO 762
LENGTH: 245
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (110)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-762

Query Match 56.1%; Score 32; DB 5; Length 245;
Best Local Similarity 45.5%; Pred. No. 80;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSSEYQILTR 11
|||:|:~|:|:
DB 168 TSTDQVLGR 178

RESULT 6
PCT-US01-11988-763
Sequence 763, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patentlin Ver. 2.1
SEQ ID NO 763
LENGTH: 290
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-763

Query Match 56.1%; Score 32; DB 1; Length 290;
Best Local Similarity 45.5%; Pred. No. 96;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 TSSEYOILTAAR 11
Db 168 TSSDHQVLSGK 178

RESULT 7

US-09-833-245-763
; Sequence 763, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 763
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-763

Query Match
Best Local Similarity 56.1%; Score 32; DB 5; Length 290;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 TSSEYOILTAAR 11
Db 168 TSSDHQVLSGK 178

RESULT 8

US-09-739-449-11547
; Sequence 11547, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 11547
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11547

Query Match
Best Local Similarity 56.1%; Score 32; DB 5; Length 344;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTAAR 12
Db 179 TALEFTILTAAR 190

RESULT 9

US-60-248-823-102
; Sequence 102, Application US/60248823
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen

TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL000949
; CURRENT APPLICATION NUMBER: US/60/248,823
; CURRENT FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 715
; TYPE: PRT
; ORGANISM: HUMAN
US-60-248-823-102

Query Match
Best Local Similarity 56.1%; Score 32; DB 6; Length 715;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 TSSEYOILTAAR 11
Db 368 TSTDQVLAAR 378

RESULT 10

US-09-739-449-8297
; Sequence 8297, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 8297
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-8297

Query Match
Best Local Similarity 54.4%; Score 31; DB 5; Length 240;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TSSEYOIL 8
Db 149 TDSYEIL 156

RESULT 11

PCT-US01-11988-133
; Sequence 133, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 359
; TYPE: PRT

ORGANISM: Homo sapiens
PCT-US01-11988-133

Query Match
Best Local Similarity 54.4%; Score 31; DB 1; Length 359;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSEYOILTA 10
Db 222 SSEYOILTA 230

RESULT 12
US-09-833-245-133
Sequence 133; Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PFS46PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 133
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-133

Query Match
Best Local Similarity 54.4%; Score 31; DB 5; Length 359;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSEYOILTA 10
Db 222 SSEYOILTA 230

RESULT 13
PCT-US01-04098A-1378
Sequence 1378; Application PC/TUS0104098A
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: PCT/US01/04098A
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03

NUMBER OF SEQ ID NOS: 3960

SOFTWARE: Custom
SEQ ID NO 1378
LENGTH: 790
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-04098A-1378

Query Match
Best Local Similarity 54.4%; Score 31; DB 1; Length 790;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSEYOILTA 9
Db 447 STEYOVT 454

RESULT 14
US-09-193-562D-34
Sequence 34; Application US/09193562D
GENERAL INFORMATION:
APPLICANT: Pauli, Benedicht U.
TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
FILE REFERENCE: 18617.0052
CURRENT APPLICATION NUMBER: US/09/193,562D
CURRENT FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: US/60/065,922
NUMBER OF SEQ ID NOS: 47
SEQ ID NO 34
LENGTH: 902
TYPE: PRT
ORGANISM: Mus musculus
US-09-193-562D-34

Query Match
Best Local Similarity 54.4%; Score 31; DB 5; Length 902;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSSEYOILTA 10
Db 366 SSSEYOILTA 375

RESULT 15
PCT-US01-04098A-1753
Sequence 1753; Application PC/TUS0104098A
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: PCT/US01/04098A
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO: 1753
LENGTH: 4544
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-04098A-1753

Query Match 54.4%; Score 31; DB 1; Length 4544;
Best Local Similarity 62.5%;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSEQOILT 9
Db 447 STEYQVVT 454

Search completed: June 4, 2001, 12:21:29
Job time: 591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:20:03 ; Search time 260.2 seconds
(without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-26
Perfect score: 64
Sequence: 1 EDGSLWCEAAT 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues 1009251

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents: AA.Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	89.1	96	23	US-60-160-209-4252
2	57	89.1	102	16	US-60-160-203-4414
3	57	89.1	261	13	US-09-245-764-7
4	57	89.1	374	21	US-09-702-021-14
5	57	89.1	399	1	PCT-US00-05882-1128
6	51	79.7	81	23	US-60-170-373-3410
7	51	79.7	81	23	US-60-177-571-3365
8	50	78.1	93	18	US-09-474-434-1676
9	50	78.1	93	23	US-60-173-463-626
10	50	78.1	93	23	US-60-173-469-1887

11	50	78.1	93	23	US-60-173-686-1676	Sequence 1676, Ap
12	50	78.1	93	23	US-60-186-282-885	Sequence 885, Ap
13	50	78.1	95	18	US-09-474-434-1311	Sequence 1311, Ap
14	50	78.1	95	23	US-60-173-469-1481	Sequence 1481, Ap
15	50	78.1	95	23	US-60-173-686-1311	Sequence 1311, Ap
16	50	78.1	95	23	US-60-186-280-495	Sequence 495, Ap
17	50	78.1	95	23	US-60-186-282-542	Sequence 542, Ap
18	50	78.1	95	23	US-60-186-282-542	Sequence 543, Ap
19	50	78.1	95	23	US-60-186-656-1326	Sequence 1326, Ap
20	50	78.1	95	23	US-60-186-656-1327	Sequence 1327, Ap
21	50	78.1	95	23	US-60-190-000-279	Sequence 279, Ap
22	50	78.1	95	23	US-60-190-000-280	Sequence 280, Ap
23	50	78.1	141	23	US-60-207-359-150	Sequence 150, Ap
24	48	75.0	58	23	US-60-160-209-3193	Sequence 3193, Ap
25	48	75.0	359	13	US-08-951-829-14	Sequence 14, Ap
26	48	75.0	359	14	US-09-040-714-14	Sequence 14, Ap
27	48	75.0	555	13	US-08-971-635-20	Sequence 20, Ap
28	48	75.0	555	13	US-08-971-635B-19	Sequence 19, Ap
29	48	75.0	555	14	US-09-073-363B-41	Sequence 41, Ap
30	48	75.0	555	15	US-09-173-151A-32	Sequence 32, Ap
31	48	75.0	605	1	PCT-US97-20201-8	Sequence 8, Ap
32	48	75.0	605	16	US-09-283-503A-8	Sequence 8, Ap
33	48	75.0	605	16	US-09-283-503B-8	Sequence 8, Ap
34	48	75.0	1010	6	US-08-258-022-4	Sequence 4, Ap
35	48	75.0	1010	10	US-08-618-953-4	Sequence 4, Ap
36	48	75.0	1020	6	US-08-258-022-3	Sequence 3, Ap
37	48	75.0	1020	10	US-08-618-953-3	Sequence 3, Ap
38	48	75.0	1125	10	US-08-687-727B-5	Sequence 5, Ap
39	48	75.0	1125	15	US-09-143-308-5	Sequence 5, Ap
40	48	75.0	1242	10	US-08-687-727B-4	Sequence 4, Ap
41	48	75.0	1242	15	US-09-143-308-4	Sequence 4, Ap
42	47	73.4	75	23	US-60-170-373-3220	Sequence 3220, Ap
43	47	73.4	75	23	US-60-170-374-2865	Sequence 2865, Ap
44	47	73.4	103	21	US-09-784-748-11	Sequence 11, Ap
45	47	73.4	105	23	US-60-196-710-6936	Sequence 6936, Ap

ALIGNMENTS

RESULT 1

US-60-160-209-4252

Sequence 4252, Application US/60160209

GENERAL INFORMATION:

APPLICANT: BONAZZI, VIVIAN

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES

FILE REFERENCE: C1000113

CURRENT APPLICATION NUMBER: US/60/160,209

CURRENT FILING DATE: 1999-10-19

NUMBER OF SEQ ID NOS: 4646

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 4252

LENGTH: 96

TYPE: PRT

ORGANISM: HUMAN

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(96)

OTHER INFORMATION: Xaa - Any Amino Acid

US-60-160-209-4252

Query Match 89.1%; Score 57; DB 23; Length 96;
Best Local Similarity: 91.7%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 EDGSLWCEAAT 12
|||||
Db 67 EDGSLWCEAAT 78

RESULT 2
US-60-160-203-4414
; Sequence 4414, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4414
; LENGTH: 102
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-4414

Query Match
Best Local Similarity 89.1%; Score 57; DB 23; Length 102;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGGLYCEAAT 12
DB 75 EDGGLYCEAAT 86

RESULT 3
US-09-245-764-7
; Sequence 7, Application US/09245764
; GENERAL INFORMATION:
; APPLICANT: Hogarth, P. Mark
; APPLICANT: Powell, Maree S.
; APPLICANT: McKenzie, Ian F.C.
; APPLICANT: Maxwell, Kelly F.
; APPLICANT: Garrett, Thomas P.J.
; APPLICANT: Epa, Vidana
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES AND MODELS OF FC RECEPTORS
; FILE REFERENCE: 4102-4
; CURRENT APPLICATION NUMBER: US/09/245,764
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/099,994
; EARLIER FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: 60/073,972
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-245-764-7

Query Match
Best Local Similarity 89.1%; Score 57; DB 16; Length 261;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGGLYCEAAT 12
DB 234 EDGGLYCEAAT 245

RESULT 4
US-09-702-021-14
; Sequence 14, Application US/09702021
; GENERAL INFORMATION:
; APPLICANT: C. Fraser
; TITLE OF INVENTION: FAST MOLECULES AND USES THEREOF
; FILE REFERENCE: 7853-217

; CURRENT APPLICATION NUMBER: US/09/702,021
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-021-14

Query Match
Best Local Similarity 89.1%; Score 57; DB 21; Length 374;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGGLYCEAAT 12
DB 253 EDGGLYCEAAT 264

RESULT 5
PCT-US00-05882-1128
; Sequence 1128, Application PC/TUS0005882
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: P106PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05882
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1128
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (208)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (349)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1128

Query Match
Best Local Similarity 89.1%; Score 57; DB 1; Length 399;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGGLYCEAAT 12
DB 278 EDGGLYCEAAT 289

RESULT 6
US-60-170-373-3410
; Sequence 3410, Application US/60170373
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: C1000149
; CURRENT APPLICATION NUMBER: US/60/170,373
; CURRENT FILING DATE: 1999-12-13
; NUMBER OF SEQ ID NOS: 4282
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3410
; LENGTH: 81

TYPE: PRT
ORGANISM: Human
US-60-170-373-3410

Query Match
Best Local Similarity 79.7%; Score 51; DB 23; Length 81;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 EDGSLYOCEAT 12
11:111111:1
56 EDAGEYOCEAST 67

RESULT 7
US-60-177-571-3365
Sequence 3365, Application US/60177571
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
FILE REFERENCE: CLO00201
CURRENT APPLICATION NUMBER: US/60/177,571
CURRENT FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 5082
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3365
LENGTH: 81
TYPE: PRT
ORGANISM: HUMAN
US-60-177-571-3365

Query Match
Best Local Similarity 79.7%; Score 51; DB 23; Length 81;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 EDGSLYOCEAT 12
11:111111:1
56 EDAGEYOCEAST 67

RESULT 8
US-09-474-434-1676
Sequence 1676, Application US/09474434
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
FILE REFERENCE: CLO00182
CURRENT APPLICATION NUMBER: US/09/474,434
CURRENT FILING DATE: 1999-12-28
NUMBER OF SEQ ID NOS: 1872
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1676
LENGTH: 93
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(93)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-474-434-1676

Query Match
Best Local Similarity 78.1%; Score 50; DB 18; Length 93;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 EDGSLYOCEA 10
1111111111

Db 66 EDGSLYOCEA 75

RESULT 9
US-60-173-463-626
Sequence 626, Application US/60173463
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN ION CHANNEL PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN ION CHANNEL PROTEINS
FILE REFERENCE: CLO00184
CURRENT APPLICATION NUMBER: US/60/173,463
CURRENT FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 626
LENGTH: 93
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(93)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-173-463-626

Query Match
Best Local Similarity 78.1%; Score 50; DB 23; Length 93;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 EDGSLYOCEA 10
1111111111
66 EDGSLYOCEA 75

RESULT 10
US-60-173-469-1887
Sequence 1887, Application US/60173469
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: CLO00185
CURRENT APPLICATION NUMBER: US/60/173,469
CURRENT FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 2120
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1887
LENGTH: 93
TYPE: PRT
ORGANISM: HUMAN
US-60-173-469-1887

Query Match
Best Local Similarity 78.1%; Score 50; DB 23; Length 93;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 EDGSLYOCEA 10
1111111111
66 EDGSLYOCEA 75

RESULT 11
US-60-173-686-1676
Sequence 1676, Application US/60173686
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
TITLE OF INVENTION: THEREOF

FILE REFERENCE: CL000182
CURRENT APPLICATION NUMBER: US/60/173,686
CURRENT FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 1872
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1676
LENGTH: 93
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(93)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-173-686-1676

Query Match
Best Local Similarity 78.1%; Score 50; DB 23; Length 93;
Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSLYQCEA 10
||| |||||
Db 66 EDGSLYQCEA 75

RESULT 12
US-60-186-282-885
Sequence 885, Application US/60186282
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
FILE REFERENCE: CL000301
CURRENT APPLICATION NUMBER: US/60/186,282
CURRENT FILING DATE: 2000-03-01
NUMBER OF SEQ ID NOS: 896
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 885
LENGTH: 93
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(93)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-186-282-885

Query Match
Best Local Similarity 78.1%; Score 50; DB 23; Length 93;
Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSLYQCEA 10
||| |||||
Db 66 EDGSLYQCEA 75

RESULT 13
US-09-474-434-1311
Sequence 1311, Application US/09474434
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
FILE REFERENCE: CL000182
CURRENT APPLICATION NUMBER: US/09/474,434
CURRENT FILING DATE: 1999-12-28
NUMBER OF SEQ ID NOS: 1872
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1311
LENGTH: 95

TYPE: PRT
ORGANISM: HUMAN
US-09-474-434-1311

Query Match
Best Local Similarity 78.1%; Score 50; DB 18; Length 95;
Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSLYQCEA 10
||| |||||
Db 70 EDGSLYQCEA 79

RESULT 14
US-60-173-469-1481
Sequence 1481, Application US/60173469
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: CL000185
CURRENT APPLICATION NUMBER: US/60/173,469
CURRENT FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 2120
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1481
LENGTH: 95
TYPE: PRT
ORGANISM: HUMAN
US-60-173-469-1481

Query Match
Best Local Similarity 78.1%; Score 50; DB 23; Length 95;
Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSLYQCEA 10
||| |||||
Db 70 EDGSLYQCEA 79

RESULT 15
US-60-173-686-1311
Sequence 1311, Application US/60173686
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
FILE REFERENCE: CL000182
CURRENT APPLICATION NUMBER: US/60/173,686
CURRENT FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 1872
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1311
LENGTH: 95
TYPE: PRT
ORGANISM: HUMAN
US-60-173-686-1311

Query Match
Best Local Similarity 78.1%; Score 50; DB 23; Length 95;
Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSLYQCEA 10
||| |||||
Db 70 EDGSLYQCEA 79

Tue Jun 5 07:09:16 2001

Search completed: June 4, 2001, 12:20:04
Job time: 517 sec

us-09-284-107-26.rapm

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:21:29 ; Search time 84.28 Seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-26

Perfect score: 64
Sequence: 1 EDSSLGLOCEAAT 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Pending Patents, AA, New: *
1: /cgn2_6/prodata1/paa/PCT_NEW_COMB.pep: *
2: /cgn2_6/prodata1/paa/US06_NEW_COMB.pep: *
3: /cgn2_6/prodata1/paa/US07_NEW_COMB.pep: *
4: /cgn2_6/prodata1/paa/US08_NEW_COMB.pep: *
5: /cgn2_6/prodata1/paa/US09_NEW_COMB.pep: *
6: /cgn2_6/prodata1/paa/US60_NEW_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	12	5	US-09-284-107-26
2	47	73.4	103	1	PCT-US01-10462-11
3	47	73.4	251	1	PCT-US01-10462-10
4	47	73.4	278	1	PCT-US01-10462-6
5	46	71.9	172	5	US-09-809-391-760
6	46	71.9	198	5	US-09-809-391-759
7	46	71.9	234	5	US-09-809-391-478
8	44	68.8	261	5	US-09-270-849B-188150
9	43	67.2	764	5	US-09-142-956B-14
10	43	67.2	789	5	US-09-232-773A-15
11	42	65.6	56	5	US-09-815-626-14
12	42	65.6	56	5	US-09-822-687-9
13	42	65.6	268	1	PCT-US01-10462-13
14	42	65.6	739	5	US-09-266-091-4
15	41	64.1	359	5	PCT-US01-11988-133
16	41	64.1	359	5	US-09-833-245-133
17	41	64.1	458	1	PCT-US00-35017A-1150
18	41	64.1	1884	6	US-06-248-505-1316
19	40	62.5	55	5	US-09-270-849B-192970
20	40	62.5	91	5	US-09-270-849B-183931
21	40	62.5	548	1	PCT-US01-09326-47
22	40	62.5	548	5	US-09-819-136-2
23	40	62.5	1745	1	PCT-US01-11988-2222
24	40	62.5	1745	5	US-09-833-245-2222
25	39	60.9	170	5	US-09-814-950-4
26	39	60.9	172	5	US-09-814-950-2
27	39	60.9	264	5	US-09-270-849B-189682

28	39	60.9	574	5	US-09-815-108-7	Sequence 7, Appl
29	39	60.9	780	5	US-09-232-773A-14	Sequence 14, Appl
30	39	60.9	1338	5	US-09-426-371-3	Sequence 3, Appl
31	38	59.4	34	5	US-09-565-522B-5	Sequence 5, Appl
32	38	59.4	86	5	US-09-270-849B-190337	Sequence 190337,
33	38	59.4	148	5	US-09-270-849B-182954	Sequence 182954,
34	38	59.4	150	5	US-09-270-849B-186137	Sequence 186137,
35	38	59.4	175	5	US-09-270-849B-186367	Sequence 186367,
36	38	59.4	249	1	PCT-US01-06769-13	Sequence 13, Appl
37	38	59.4	264	1	PCT-US01-10462-14	Sequence 14, Appl
38	38	59.4	342	5	US-09-815-108-20	Sequence 20, Appl
39	38	59.4	448	5	US-09-815-108-6	Sequence 6, Appl
40	38	59.4	472	5	US-09-815-108-5	Sequence 5, Appl
41	38	59.4	504	5	US-09-815-108-8	Sequence 8, Appl
42	38	59.4	504	5	US-09-815-108-15	Sequence 15, Appl
43	38	59.4	504	5	US-09-815-108-17	Sequence 17, Appl
44	38	59.4	504	5	US-09-815-108-19	Sequence 19, Appl
45	38	59.4	509	5	US-09-815-108-3	Sequence 3, Appl

ALIGNMENTS

```
RESULT 1
US-09-284-107-26
Sequence 26, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
APPLICANT: de Kruijf, Cornelis Adriaan John
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: AND PROTEINS HAVING SPECIFIC AFFINITY FOR A TARGET
FILE REFERENCE: 313632000600
CURRENT APPLICATION NUMBER: US/09/284,107
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
PRIOR FILING DATE: 1996-10-08
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
US-09-284-107-26
Query Match 100.0%; Score 64; DB 5; Length 12;
Best local similarity 100.0%; Pred. No. 1.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDSSLGLOCEAAT 12
DB 1 EDSSLGLOCEAAT 12
RESULT 2
PCT-US01-10462-11
Sequence 11, Application PC/TUS0110462
GENERAL INFORMATION:
APPLICANT: HYSEQ, INC.
APPLICANT: Yamazaki, Victoria
APPLICANT: Asundi, Vinod
APPLICANT: Liu, Chenchua
APPLICANT: Tang, Y. Tom
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO INSULIN-LIKE GROWTH FACTOR B
FILE REFERENCE: 21272-051
CURRENT APPLICATION NUMBER: PCT/US01/10462
```

CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/784,748
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
LENGTH: 103
TYPE: PRT
ORGANISM: homo sapiens
PCT-US01-10462-11

Query Match 73.4%; Score 47; DB 1; Length 103;
Best Local Similarity 72.7%; Pred. No. 0.18;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDGGLYQCEAA 11
|||:|||||
Db 91 EDEGVYQCHAA 101

RESULT 3
PCT-US01-10462-10
Sequence 10, Application PC/TUS0110462
GENERAL INFORMATION:
APPLICANT: HYSEO, INC.
APPLICANT: Yamazaki, Victoria
APPLICANT: Asundi, Vinod
APPLICANT: Liu, Chenghua
APPLICANT: Tang, Y. Tom
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO INSULIN-LIKE GROWTH FACTOR BIND
FILE REFERENCE: 21272-051
CURRENT APPLICATION NUMBER: PCT/US01/10462
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/784,748
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 251
TYPE: PRT
ORGANISM: homo sapiens
PCT-US01-10462-10

Query Match 73.4%; Score 47; DB 1; Length 251;
Best Local Similarity 72.7%; Pred. No. 0.46;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDGGLYQCEAA 11
|||:|||||
Db 209 EDEGVYQCHAA 219

RESULT 4
PCT-US01-10462-6
Sequence 6, Application PC/TUS0110462
GENERAL INFORMATION:
APPLICANT: HYSEO, INC.
APPLICANT: Yamazaki, Victoria
APPLICANT: Asundi, Vinod
APPLICANT: Liu, Chenghua
APPLICANT: Tang, Y. Tom

APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO INSULIN-LIKE GROWTH FACTOR B
FILE REFERENCE: 21272-051
CURRENT APPLICATION NUMBER: PCT/US01/10462
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/784,748
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 278
TYPE: PRT
ORGANISM: homo sapiens
PCT-US01-10462-6

Query Match 73.4%; Score 47; DB 1; Length 278;
Best Local Similarity 72.7%; Pred. No. 0.51;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDGGLYQCEAA 11
|||:|||||
Db 236 EDEGVYQCHAA 246

RESULT 5
US-09-809-391-760
Sequence 760, Application US/09809391
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P2
CURRENT APPLICATION NUMBER: US/09/809,391
CURRENT FILING DATE: 2001-03-16
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 761
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 760
LENGTH: 172
TYPE: PRT
ORGANISM: Homo sapiens
US-09-809-391-760

Query Match 71.9%; Score 46; DB 5; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGGLYQC 8
|||:|||||
Db 80 EDEGLYQC 87

RESULT 6
US-09-809-391-759
Sequence 759, Application US/09809391
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P2
CURRENT APPLICATION NUMBER: US/09/809,391
CURRENT FILING DATE: 2001-03-16
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 761
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 759
LENGTH: 198
TYPE: PRT

ORGANISM: Homo sapiens
US-09-809-391-759

Query Match 71.9%; Score 46; DB 5; Length 198;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGGLYOC 8
Db 106 EDGGLYOC 113

RESULT 7
US-09-809-391-478
Sequence 478, Application US/09809391
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P2
CURRENT APPLICATION NUMBER: US/09/809,391
CURRENT FILING DATE: 2001-03-16
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 761
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 478
LENGTH: 234
TYPE: PRT
ORGANISM: Homo sapiens
US-09-809-391-478

Query Match 71.9%; Score 46; DB 5; Length 234;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGGLYOC 8
Db 106 EDGGLYOC 113

RESULT 8
US-09-270-849B-188150
Sequence 188150, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 188150
LENGTH: 261
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-188150

Query Match 68.8%; Score 44; DB 5; Length 261;
Best Local Similarity 54.5%; Pred. No. 1.7;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 DSGLYOCEAT 12
Db 68 DSGLYOCEAT 78

RESULT 9
US-09-142-956B-14

Sequence 14, Application US/09142956B
GENERAL INFORMATION:
APPLICANT: Niwa, Mikio
APPLICANT: Okamoto, Masaji
APPLICANT: Matsumoto, Tomoe
APPLICANT: Segawa, Toshiaki
TITLE OF INVENTION: VEGF-BINDING POLYPEPTIDES
FILE REFERENCE: 06501-021001
CURRENT APPLICATION NUMBER: US/09/142,956B
CURRENT FILING DATE: 1999-03-04
PRIOR APPLICATION NUMBER: PCT/JP98/00140
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: JP 9/19706
PRIOR FILING DATE: 1997-01-17
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 764
TYPE: PRT
ORGANISM: Homo sapiens
US-09-142-956B-14

Query Match 67.2%; Score 43; DB 5; Length 764;
Best Local Similarity 58.3%; Pred. No. 7.7;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGGLYOC 12
Db 730 EDGGLYOC 741

RESULT 10
US-09-232-773A-15
Sequence 15, Application US/09232773A
GENERAL INFORMATION:
APPLICANT: Kendall, Richard L.
APPLICANT: Thomas, Kenneth A., Jr.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
GROWTH FACTOR
FILE REFERENCE: 18888DB
CURRENT APPLICATION NUMBER: US/09/232,773A
CURRENT FILING DATE: 1999-07-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 789
TYPE: PRT
ORGANISM: Homo sapiens
US-09-232-773A-15

Query Match 67.2%; Score 43; DB 5; Length 789;
Best Local Similarity 58.3%; Pred. No. 8;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGGLYOC 12
Db 730 EDGGLYOC 741

RESULT 11
US-09-815-626-14
Sequence 14, Application US/09815626
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria Alexandra
TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT
FAMILY MEMBER AND USES THEREOF
FILE REFERENCE: 10448-031001
CURRENT APPLICATION NUMBER: US/09/815,626
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,863
PRIOR FILING DATE: 2000-03-24

NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 56
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-09-815-626-14

Query Match
Best Local Similarity 65.6%; Score 42; DB 5; Length 56;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDGSLYXCEA 10
Db 42 EDGSLYTCVA 51

RESULT 12
US-09-822-687-9
Sequence 9, Application US/09822687
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria Alexandra
TITLE OF INVENTION: 31939, A NOVEL HUMAN LEUCINE-RICH REPEAT
FILE REFERENCE: 10448-038001
CURRENT APPLICATION NUMBER: US/09/822,687
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,919
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 56
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-09-822-687-9

Query Match
Best Local Similarity 65.6%; Score 42; DB 5; Length 56;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDGSLYXCEA 10
Db 42 EDGSLYTCVA 51

RESULT 13
PCT-US01-10462-13
Sequence 13, Application PC/TUS0110462
GENERAL INFORMATION:
APPLICANT: HISEQ, INC.
APPLICANT: Yamazaki, Victoria
APPLICANT: Asundi, Vinod
APPLICANT: Liu, Chenghua
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO INSULIN-LIKE GROWTH FACTOR BINT
FILE REFERENCE: 2172-051
CURRENT APPLICATION NUMBER: PCT/US01/10462
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/784,748
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31

NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.0
SEQ ID NO 13
LENGTH: 268
TYPE: PRT
ORGANISM: mus musculus
PCT-US01-10462-13

Query Match
Best Local Similarity 65.6%; Score 42; DB 1; Length 268;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDGSLYXCEA 11
Db 228 EDEGVYHCHAA 238

RESULT 14
US-09-266-091-4
Sequence 4, Application US/09266091
GENERAL INFORMATION:
APPLICANT: Klimuk, Sandra K
APPLICANT: Semple, Sean C
APPLICANT: Scherrer, Peter
APPLICANT: Hope, Michael J.
TITLE OF INVENTION: ENHANCED EFFICACY OF LIPOSOMAL ANTISENSE THERAPY
FILE REFERENCE: ISPH-0342
CURRENT APPLICATION NUMBER: US/09/266,091
CURRENT FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 08/657,753
PRIOR FILING DATE: 1996-05-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 739
TYPE: PRT
ORGANISM: Homo sapiens
US-09-266-091-4

Query Match
Best Local Similarity 65.6%; Score 42; DB 5; Length 739;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSLYXCE 9
Db 284 EDGSLYVCE 292

RESULT 15
PCT-US01-11988-133
Sequence 133, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: P5546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 133
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-133

Query Match 64.18; Score 41; DB 1; Length 359;
Best Local Similarity 66.7%; Pred. No. 8.3;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EDGSLYCEAAT 12
: ||| |||||
Db 233 DHGSGYWCFAAT 244

Search completed: June 4, 2001, 12:21:29
Job time: 591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:20:04 ; Search time 260.2 Seconds
(Without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-27

Perfect score: 61

Sequence: 1 EDGNVTKRSPKL 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCNUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description
	1	61	100.0	96	23	US-60-160-209-4252	Sequence 4252, Ap
	2	61	100.0	102	23	US-60-160-203-4414	Sequence 4414, Ap
	3	61	100.0	261	16	US-09-245-764-7	Sequence 7, Appl
	4	61	100.0	374	21	US-09-702-021-14	Sequence 14, Appl
	5	61	100.0	399	1	PCT-US00-05882-1128	Sequence 1128, Ap
	6	50	82.0	91	7	US-08-332-562-94	Sequence 94, Appl
	7	50	82.0	91	8	US-08-462-973-94	Sequence 94, Appl
	8	50	82.0	410	3	US-07-861-895-25	Sequence 25, Appl
	9	42	68.9	99	23	US-60-191-637-27859	Sequence 27859, A
	10	42	68.9	99	23	US-60-191-681-22468	Sequence 22468, A

11	42	68.9	299	23	US-60-167-217-7942	Sequence 7942, Ap
12	42	68.9	407	17	US-09-347-801-23	Sequence 23, Appl
13	39	63.9	368	16	US-09-248-796-16492	Sequence 16492, A
14	38	62.3	490	16	US-09-248-796-16167	Sequence 16167, A
15	37.5	61.5	166	18	US-09-450-969-7543	Sequence 7543, Ap
16	37	60.7	124	21	US-09-733-089-16965	Sequence 16965, A
17	37	60.7	143	15	US-09-134-000-4294	Sequence 4294, Ap
18	37	60.7	168	23	US-60-167-217-7877	Sequence 7877, Ap
19	37	60.7	397	18	US-09-489-039A-13098	Sequence 13098, A
20	37	60.7	1087	18	US-09-417-507-42680	Sequence 42680, A
21	37	60.7	1354	20	US-09-619-049-411	Sequence 411, Appl
22	37	60.7	1354	23	US-60-167-217-7900	Sequence 7900, Ap
23	37	60.7	1354	23	US-60-171-627-660	Sequence 660, Appl
24	37	60.7	1354	23	US-60-173-464-6310	Sequence 6310, Appl
25	37	60.7	1354	23	US-60-191-637-7800	Sequence 7800, Ap
26	37	60.7	1354	23	US-60-191-681-6078	Sequence 6078, Ap
27	37	60.7	1907	23	US-60-173-464-28304	Sequence 28304, A
28	37	60.7	1908	23	US-60-191-637-36817	Sequence 36817, A
29	37	60.7	1908	23	US-60-191-681-28768	Sequence 28768, A
30	37	60.7	1970	23	US-60-167-217-22133	Sequence 22133, A
31	37	60.7	1970	23	US-60-173-464-17967	Sequence 17967, A
32	36.5	59.8	366	15	US-09-107-532-4292	Sequence 4292, Ap
33	36	59.0	56	21	US-09-733-089-15889	Sequence 15889, A
34	36	59.0	132	1	PCT-US97-14436-634	Sequence 634, Appl
35	36	59.0	132	13	US-08-911-503-634	Sequence 634, Appl
36	36	59.0	132	1	PCT-US01-01326-137	Sequence 137, Appl
37	36	59.0	181	1	PCT-US01-01349-834	Sequence 834, Appl
38	36	59.0	181	16	US-09-270-767-35780	Sequence 35780, A
39	36	59.0	189	16	US-09-270-767-50997	Sequence 50997, A
40	36	59.0	213	1	PCT-US96-15098A-11	Sequence 11, Appl
41	36	59.0	217	1	PCT-US01-01349-664	Sequence 664, Appl
42	36	59.0	260	18	US-09-497-191-861	Sequence 861, Appl
43	36	59.0	286	18	US-09-450-969-5322	Sequence 5322, Appl
44	36	59.0	358	16	US-09-206-647-1	Sequence 1, Appl
45	36	59.0				

ALIGNMENTS

RESULT 1
US-60-160-209-4252
Sequence 4252, Application US/60160209
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIEN
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
FILE REFERENCE: C1000113
CURRENT APPLICATION NUMBER: US/60/160,209
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 4646
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4252
LENGTH: 96
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(96)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-209-4252

Query Match 100.0%; Score 61; DB 23; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 EDGNVTKRSPKL 12
|||||
79 EDGNVTKRSPKL 90

RESULT 2
US-60-160-203-4414
; Sequence 4414, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4414
; LENGTH: 102
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-4414

Query Match
Best Local Similarity 100.0%; Score 61; DB 23; Length 102;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRSPSEL 12
Db 87 EDGNVLRSPSEL 98

RESULT 3
US-09-245-764-7
; Sequence 7, Application US/09245764
; GENERAL INFORMATION:
; APPLICANT: HOGARTH, P. Mark
; APPLICANT: POWELL, MAREE S.
; APPLICANT: MCKENZIE, IAN F.C.
; APPLICANT: MAXWELL, KELLY F.
; APPLICANT: GARRETT, THOMAS P.J.
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES AND MODELS OF FC RECEPTORS
; FILE REFERENCE: 4102-4
; CURRENT APPLICATION NUMBER: US/09/245,764
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/099,994
; EARLIER FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: 60/073,972
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-245-764-7

Query Match
Best Local Similarity 100.0%; Score 61; DB 16; Length 261;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRSPSEL 12
Db 246 EDGNVLRSPSEL 257

RESULT 4
US-09-702-021-14
; Sequence 14, Application US/09702021
; GENERAL INFORMATION:
; APPLICANT: C. FRASER
; TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF
; FILE REFERENCE: 7853-217

CURRENT APPLICATION NUMBER: US/09/702,021
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-021-14

Query Match
Best Local Similarity 100.0%; Score 61; DB 21; Length 374;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRSPSEL 12
Db 265 EDGNVLRSPSEL 276

RESULT 5
PCT-US00-05882-1128
; Sequence 1128, Application PC/TUS0005882
; GENERAL INFORMATION:
; APPLICANT: CRAIG ROSEN,
; APPLICANT: STEVE RUBEN,
; TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: PA106PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05882
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1128
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (208)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (349)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1128

Query Match
Best Local Similarity 100.0%; Score 61; DB 1; Length 399;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRSPSEL 12
Db 290 EDGNVLRSPSEL 301

RESULT 6
US-08-332-562-94
; Sequence 94, Application US/08332562
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN F.C.
; APPLICANT: HOGARTH, MARK P.
; APPLICANT: HIBBS, MARGARET L.
; APPLICANT: SCOTT, BERNADETTE M.
; APPLICANT: BONADONNA, LISA
; APPLICANT: HULETT, MARK D.
; TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W., Suite 500

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562-94

Query Match 82.0%; Score 50; DB 7; Length 91;
Best Local Similarity 83.3%; Pred. No. 0.052;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGNYLKRSPEL 12
DB 77 EDSVVKRSPEL 88

RESULT 7
US-08-462-973-94
Sequence 94, Application US/08462973
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F. C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,973
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/332,562
FILING DATE: 31-OCT-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH8232/94
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU87/00159
FILING DATE: 29-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH6166
FILING DATE: 26-MAY-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-462-973-94

Query Match 82.0%; Score 50; DB 8; Length 91;
Best Local Similarity 83.3%; Pred. No. 0.052;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGNYLKRSPEL 12
DB 77 EDSVVKRSPEL 88

RESULT 8
US-07-861-895-25
Sequence 25, Application US/07861895
GENERAL INFORMATION:
APPLICANT: HOGARTH, Phillip M.
APPLICANT: HULETT, Mark D.
APPLICANT: IERINO, Francesco L.
APPLICANT: MCKENZIE, Ian F. C.
APPLICANT: OSMAN, Nafin
TITLE OF INVENTION: HYBRID FC RECEPTOR MOLECULES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/861,895
FILING DATE: 19920624
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00513
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7045
FILING DATE: 25-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 17227/111
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 410 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein
US-07-861-895-25

Query Match. 82.0%; Score 50; DB 3; Length 410;
Best Local Similarity 83.3%; Pred. No. 0.36;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGNVLRKSPDL 12
||:|||||
Db 280 EDSVLRKSPDL 291

RESULT 9
US-60-191-637-27859

Sequence 27859, Application US/60191637

GENERAL INFORMATION:

APPLICANT: Venter, J. Craig

TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING

TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: CL000392

CURRENT APPLICATION NUMBER: US/60/191,637

CURRENT FILING DATE: 2000-03-23

NUMBER OF SEQ ID NOS: 42660

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 27859

LENGTH: 99

TYPE: PRT

ORGANISM: DROSOPHILA

US-60-191-637-27859

Query Match 68.9%; Score 42; DB 23; Length 99;
Best Local Similarity 54.5%; Pred. No. 2.2;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGNVLRKSPDL 11
||:|||||
Db 36 DDGNVLRKSPDL 46

RESULT 10

US-60-191-681-22468

Sequence 22468, Application US/60191681

GENERAL INFORMATION:

APPLICANT: Li, Peter, W.D.

TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: CL000390

CURRENT APPLICATION NUMBER: US/60/191,681

CURRENT FILING DATE: 2000-03-23

NUMBER OF SEQ ID NOS: 30973

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 22468

LENGTH: 99

TYPE: PRT

ORGANISM: DROSOPHILA

US-60-191-681-22468

Query Match 68.9%; Score 42; DB 23; Length 99;

Best Local Similarity 54.5%; Pred. No. 2.2;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGNVLRKSPDL 11
||:|||||
Db 36 DDGNVLRKSPDL 46

RESULT 11

US-60-167-217-7942

Sequence 7942, Application US/60167217

GENERAL INFORMATION:

APPLICANT: Li, Peter W. D.

TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: CL000152

CURRENT APPLICATION NUMBER: US/60/167,217

CURRENT FILING DATE: 1999-11-24

NUMBER OF SEQ ID NOS: 23195

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 7942

LENGTH: 299

TYPE: PRT

ORGANISM: Drosophila

US-60-167-217-7942

Query Match 68.9%; Score 42; DB 23; Length 299;
Best Local Similarity 54.5%; Pred. No. 8.9;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGNVLRKSPDL 11
||:|||||
Db 236 DDGNVLRKSPDL 246

RESULT 12

US-09-347-801-23

Sequence 23, Application US/09347801

GENERAL INFORMATION:

APPLICANT: Allen, Steve

APPLICANT: Lee, Jian Ming

TITLE OF INVENTION: Plant Protein Kinases

FILE REFERENCE: BB-11771

CURRENT APPLICATION NUMBER: US/09/347,801

CURRENT FILING DATE: 1999-07-02

EARLIER APPLICATION NUMBER: 60/092,438

EARLIER FILING DATE: July 10, 1998

NUMBER OF SEQ ID NOS: 23

SOFTWARE: Microsoft Office 97

SEQ ID NO 23

LENGTH: 407

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-09-347-801-23

Query Match 68.9%; Score 42; DB 17; Length 407;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDGNVLRKSPDL 12
||:|||||
Db 21 DDGNVLRKSPDL 31

RESULT 13

US-09-248-796-16492

Sequence 16492, Application US/09248796

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI

TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248.796

CURRENT FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 28206

SEQ ID NO 16492

LENGTH: 368

TYPE: PRT

ORGANISM: Candida albicans

US-09-248-796-16492

Query Match

Best Local Similarity 63.9%; Score 39; DB 16; Length 368;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRSPDL 12

Db 167 EIGNVLRSPDL 178

RESULT 14

US-09-248-796-16167

Sequence 16167, Application US/09248796

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248.796

CURRENT FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 28206

SEQ ID NO 16167

LENGTH: 490

TYPE: PRT

ORGANISM: Candida albicans

US-09-248-796-16167

Query Match

Best Local Similarity 62.3%; Score 38; DB 16; Length 490;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRSPDL 12

Db 330 KGNVLRSPDL 341

RESULT 15

US-09-450-969-7543

Sequence 7543, Application US/09450969

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

FILE REFERENCE: PAT99-09A

CURRENT APPLICATION NUMBER: US/09/450.969

CURRENT FILING DATE: 1999-11-29

NUMBER OF SEQ ID NOS: 7544

SEQ ID NO 7543

LENGTH: 166

TYPE: PRT

ORGANISM: S.epidermidis

US-09-450-969-7543

Query Match

Best Local Similarity 61.5%; Score 37.5; DB 18; Length 166;

Matches 8; Conservative 4; Mismatches 0; Indels 7; Gaps 1;

OY 1 EDGNVLRSPDL 12

Db 109 EGNVLRSPDL 127

Search completed: June 4, 2001, 12:20:04
Job time: 517 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein . protein search, using sw model

Run on: June 4, 2001, 12:21:29 ; Search time 84.28 Seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-27

Perfect score: 61

Sequence: 1 EDGNVLRKSPDL 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptoddata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptoddata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptoddata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptoddata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptoddata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptoddata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	12	5	US-09-284-107-27
2	36	59.0	124	5	US-09-739-449-11283
3	36	59.0	189	5	US-09-270-8498-183641
4	36	59.0	359	1	PCT-US01-11988-133
5	36	59.0	359	5	US-09-833-245-133
6	35	57.4	133	5	US-09-270-8498-192159
7	35	57.4	470	5	US-09-270-8498-191599
8	35	57.4	497	5	US-09-549-066-37
9	35	57.4	503	5	US-09-549-066-36
10	34	55.7	94	1	PCT-US01-01312-1066
11	34	55.7	216	5	US-09-739-449-12341
12	34	55.7	223	1	PCT-US01-04098A-1261
13	34	55.7	232	1	PCT-US01-04098A-3229
14	34	55.7	801	1	PCT-US01-11797-29
15	33	54.1	184	5	US-09-739-449-11837
16	33	54.1	202	5	US-09-739-449-11789
17	33	54.1	365	1	PCT-US01-04098A-1282
18	33	54.1	381	1	PCT-US01-04098A-3250
19	33	54.1	382	1	PCT-US01-04098A-1251
20	33	54.1	480	5	US-09-672-459-4
21	33	54.1	483	5	US-09-291-023A-16
22	33	54.1	503	5	US-09-549-066-2
23	33	54.1	503	5	US-09-549-066-35
24	33	54.1	514	5	US-09-327-563B-4
25	33	54.1	539	1	PCT-US01-01332-762
26	33	54.1	633	5	US-09-325-430B-5
27	33	54.1	633	5	US-09-325-430B-5

28	33	54.1	665	5	US-09-816-494-2	Sequence 2, Appl1
29	33	54.1	672	1	PCT-US00-35017A-1259	Sequence 1559, Ap
30	33	54.1	1359	6	US-60-248-823-85	Sequence 85, Appl
31	32	52.5	130	5	US-09-270-8498-185071	Sequence 185071,
32	32	52.5	175	5	US-09-817-427-595	Sequence 395, App
33	32	52.5	231	5	US-09-739-449-8764	Sequence 8764, Ap
34	32	52.5	344	5	US-09-739-449-10210	Sequence 10210, A
35	32	52.5	415	5	US-09-826-212-6	Sequence 14, Appl
36	32	52.5	425	5	US-09-486-734A-14	Sequence 1027, Ap
37	32	52.5	501	5	US-09-640-211A-1027	Sequence 3, Appl1
38	32	52.5	623	5	US-09-825-414-3	Sequence 13, Appl
39	32	52.5	756	5	US-09-707-468A-13	Sequence 2, Appl1
40	32	52.5	906	5	US-09-459-715-2	Sequence 1, Appl1
41	32	52.5	1033	5	US-09-834-309-1	Sequence 723, App
42	32	52.5	1242	6	US-60-248-505-723	Sequence 1162, Ap
43	32	52.5	1247	6	US-60-248-505-1162	Sequence 529, App
44	31	50.8	44	4	US-08-467-344A-529	Sequence 874, App
45	31	50.8	204	1	PCT-US01-01332-874	

ALIGNMENTS

```
RESULT 1
US-09-284-107-27
Sequence 27, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: AND PROTEINS HAVING SPECIFIC AFFINITY FOR A TARGET
FILE REFERENCE: 313632000600
CURRENT APPLICATION NUMBER: US/09/284,107
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
PRIOR FILING DATE: 1996-10-08
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 27
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
US-09-284-107-27

Query Match
Best Local Similarity 100.0%; Score 61; DB 5; Length 12;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRKSPDL 12
DB 1 EDGNVLRKSPDL 12

RESULT 2
US-09-739-449-11283
Sequence 11283, Application US/09739449
GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
FILE REFERENCE: 38-101154901C
CURRENT APPLICATION NUMBER: US/09/739,449
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 11283
```

LENGTH: 124
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11283

Query Match 59.0%; Score 36; DB 5; Length 124;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DGNVLRSP 12
DB 99 DGRILKRGHEL 109

RESULT 3
US-09-270-849B-183641
Sequence 183641, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270, 849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 183641
LENGTH: 189
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-183641

Query Match 59.0%; Score 36; DB 5; Length 189;
Best Local Similarity 63.6%; Pred. No. 34;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGNVLRSP 11
DB 108 EDCDLERSVE 118

RESULT 4
PCT-US01-11988-133
Sequence 133, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 133
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-133

Query Match 59.0%; Score 36; DB 1; Length 359;
Best Local Similarity 58.3%; Pred. No. 73;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDGNVLRSP 12
DB 245 EDNQVWKOSPOL 256

RESULT 5
US-09-833-245-133
Sequence 133, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833, 245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 133
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-133

Query Match 59.0%; Score 36; DB 5; Length 359;
Best Local Similarity 58.3%; Pred. No. 73;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDGNVLRSP 12
DB 245 EDNQVWKOSPOL 256

RESULT 6
US-09-270-849B-192159
Sequence 192159, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270, 849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 192159
LENGTH: 193
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-192159

Query Match 57.4%; Score 35; DB 5; Length 193;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDGNVLRSP 10
DB 6 EDGVVLRPP 15

RESULT 7
US-09-270-849B-191599
Sequence 191599, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices

FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 191599
LENGTH: 470
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-191599

Query Match 57.4%; Score 35; DB 5; Length 470;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNTLKRPSP 10
Db 282 EDGNTLKRPSP 291

RESULT 8
US-09-549-066-37
Sequence 37, Application US/09549066
GENERAL INFORMATION:
APPLICANT: Yuan, Junying
APPLICANT: Miura, Masayuki
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/549,066
FILING DATE: 13-APRIL-2000
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/258,287B
FILING DATE: 10-JUN-1994
APPLICATION NUMBER: US 08/080,850
FILING DATE: 24-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609,3920001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: both
US-09-549-066-37

Query Match 57.4%; Score 35; DB 5; Length 497;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNTLKRPSP 12

Db 465 QGNTLKMPSP 476

RESULT 9
US-09-549-066-36
Sequence 36, Application US/09549066
GENERAL INFORMATION:
APPLICANT: Yuan, Junying
APPLICANT: Miura, Masayuki
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/549,066
FILING DATE: 13-APRIL-2000
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/258,287B
FILING DATE: 10-JUN-1994
APPLICATION NUMBER: US 08/080,850
FILING DATE: 24-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609,3920001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: both
US-09-549-066-36

Query Match 57.4%; Score 35; DB 5; Length 503;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNTLKRPSP 12
Db 471 QGNTLKMPSP 482

RESULT 10
PCT-US01-01312-1066
Sequence 1066, Application PCT/US0101312
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTA32PCT
CURRENT APPLICATION NUMBER: PCT/US01/01312
CURRENT FILING DATE: 2001-05-09
Prior application data removed - refer to PAM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1066
LENGTH: 94
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: SITE
LOCATION: (32)
OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
PCT-US01-01312-1066

Query Match 55.7%; Score 34; DB 1; Length 94;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GNVLRSPDL 12
Db 43 GNVLRSPDL 52

RESULT 11
US-09-739-449-12341

Sequence 12341, Application US/09739449
GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(13490)C
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US/09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 12341
LENGTH: 216
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-12341

Query Match 55.7%; Score 34; DB 5; Length 216;
Best Local Similarity 87.5%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GNVLRSPDL 10
Db 180 GNVLRSPDL 187

RESULT 12
PCT-US01-04098A-1261

Sequence 1261, Application PC/TUS0104098A
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: Not yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960

SOFTWARE: Custom
SEQ ID NO 1261
LENGTH: 223
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-04098A-1261

Query Match 55.7%; Score 34; DB 1; Length 223;
Best Local Similarity 47.4%; Pred. No. 99;
Matches 9; Conservative 2; Mismatches 0; Indels 8; Gaps 1;

QY 1 EDGVLKRPDL-----SPE 11
Db 109 EDGVLKRPDLSETEKMSPE 127

RESULT 13

PCT-US01-04098A-3229
Sequence 3229, Application PC/TUS0104098A
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: PCT/US01/04098A
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: Not yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 3229
LENGTH: 232
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-04098A-3229

Query Match 55.7%; Score 34; DB 1; Length 232;
Best Local Similarity 47.4%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 0; Indels 8; Gaps 1;

QY 1 EDGVLKRPDL-----SPE 11
Db 118 EDGVLKRPDLSETEKMSPE 136

RESULT 14
PCT-US01-11797-29

Sequence 29, Application PC/TUS0111797
GENERAL INFORMATION:
APPLICANT: SMITHKLINE BEECHAM CORPORATION
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP500022
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 60/196,603

PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 60/199,417
PRIOR FILING DATE: 2000-04-24
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 29
LENGTH: 801
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11797-29

Query Match 55.7% Score 34; DB 1; Length 801;
Best Local Similarity 54.5% Pred. NO. 4.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRKSP 11
Db 599 DDGHVMSCSPE 609

RESULT 15
US-09-739-449-11837
Sequence 11837, Application US/09739449
GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15490)C
CURRENT APPLICATION NUMBER: US/09/739,449
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 11837
LENGTH: 184
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11837

Query Match 54.1% Score 33; DB 5; Length 184;
Best Local Similarity 66.7% Pred. NO. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DGNVLRKSP 10
Db 133 DGLLRKSP 141

Search completed: June 4, 2001, 12:21:29
Job time: 591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:20:04 ; Search time 260.2 Seconds

(without alignments)
7.418 Million cell updates/sec

Title: US-09-284-107-28

Perfect score: 60

Sequence: 1 ELQVIGLQLPTP 12

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

1009251

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCRTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	374	21	US-09-702-021-14
2	60	100.0	399	1	PCT-US00-05882-1128
3	43	71.7	165	1	PCT-US01-03080-9
4	43	71.7	195	1	PCT-US01-03080-8
5	39	65.0	214	23	US-60-258-251-43
6	39	65.0	338	18	US-09-467-008-6
7	39	65.0	380	3	US-07-922-466-6
8	39	65.0	380	3	US-07-922-466A-6
9	39	65.0	380	7	US-08-394-442-6
10	39	65.0	380	8	US-08-416-478-6

11	39	65.0	380	8	US-08-474-988-6	Sequence 6, Appl
12	39	65.0	380	8	US-08-474-988A-6	Sequence 6, Appl
13	39	65.0	422	18	US-09-467-008-4	Sequence 7, Appl
14	39	65.0	470	7	US-08-394-442-7	Sequence 7, Appl
15	39	65.0	470	8	US-08-474-988-7	Sequence 7, Appl
16	39	65.0	470	8	US-08-474-988A-7	Sequence 7, Appl
17	39	65.0	471	7	US-08-394-442-2	Sequence 2, Appl
18	39	65.0	471	8	US-08-474-988-2	Sequence 2, Appl
19	39	65.0	471	8	US-08-474-988A-2	Sequence 2, Appl
20	39	65.0	498	3	US-07-922-466-2	Sequence 2, Appl
21	39	65.0	498	3	US-07-922-466A-2	Sequence 2, Appl
22	39	65.0	498	7	US-08-394-442-9	Sequence 9, Appl
23	39	65.0	498	8	US-08-416-478-2	Sequence 9, Appl
24	39	65.0	498	8	US-08-474-988-9	Sequence 9, Appl
25	39	65.0	498	8	US-08-474-988A-9	Sequence 9, Appl
26	38	63.3	126	23	US-60-197-873-14657	Sequence 14657, A
27	38	63.3	131	23	US-60-147-499-4839	Sequence 4839, Ap
28	38	63.3	145	23	US-60-197-873-17725	Sequence 17725, A
29	38	63.3	514	16	US-09-273-420-2	Sequence 2, Appl
30	38	63.3	515	16	US-09-273-420-6	Sequence 6, Appl
31	38	63.3	577	20	US-09-684-458-2	Sequence 2, Appl
32	38	63.3	581	16	US-09-273-420-10	Sequence 10, Appl
33	38	63.3	610	16	US-09-273-420-4	Sequence 4, Appl
34	37	61.7	264	16	US-09-201-228A-684	Sequence 684, App
35	36	60.0	73	1	PCT-US97-15872-50	Sequence 50, Appl
36	36	60.0	73	11	US-08-706-741A-50	Sequence 50, Appl
37	36	60.0	73	11	US-08-706-741A-50	Sequence 50, Appl
38	36	60.0	73	13	US-08-924-695-50	Sequence 50, Appl
39	36	60.0	127	21	US-09-733-089-3657	Sequence 3657, Ap
40	36	60.0	128	23	US-60-182-485-57	Sequence 57, Appl
41	36	60.0	128	23	US-60-182-485-58	Sequence 58, Appl
42	36	60.0	157	23	US-60-170-374-3209	Sequence 3209, Ap
43	36	60.0	158	15	US-09-107-532-4621	Sequence 4621, Ap
44	36	60.0	165	23	US-60-177-670-182	Sequence 182, App
45	36	60.0	175	23	US-60-177-670-133	Sequence 133, App

ALIGNMENTS

RESULT 1

US-09-702-021-14

Sequence 14, Application US/09702021

GENERAL INFORMATION:

APPLICANT: C. Fraser

TITLE OF INVENTION: FAIL MOLECULES AND USES THEREOF

FILE REFERENCE: 7853-217

CURRENT APPLICATION NUMBER: US/09/702,021

CURRENT FILING DATE: 2000-10-30

NUMBER OF SEQ ID NOS: 45

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 14

LENGTH: 374

TYPE: PRT

ORGANISM: Homo sapiens

US-09-702-021-14

Query Match 100.0%; Score 60; DB 21; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.073;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELQVIGLQLPTP 12

DB 277 ELQVIGLQLPTP 288

RESULT 2

PCT-US00-05882-1128

Sequence 1128, Application PC/TUS0005882

GENERAL INFORMATION:

APPLICANT: Craig Rosen,

APPLICANT: Steve Ruben

TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
FILE REFERENCE: P0106PCT
CURRENT APPLICATION NUMBER: PCT/US00/05882
CURRENT FILING DATE: 2000-03-08
EARLIER APPLICATION NUMBER: 60/124,270
EARLIER FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1128
LENGTH: 399
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: SITE
LOCATION: (208)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (349)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1128

Query Match 100.0%; Score 60; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.078;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQPTP 12
||:|||||
Db 302 ELQVIGLQPTP 313

RESULT 3
PCT-US01-03080-9
Sequence 9, Application PC/TUS0103080
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Bcl-2-like Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: P0044PCT
CURRENT APPLICATION NUMBER: PCT/US01/03080
CURRENT FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/179,487
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/180,697
PRIOR FILING DATE: 2000-02-07
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 165
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: SITE
LOCATION: (126)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (158)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-03080-9

Query Match 71.7%; Score 43; DB 1; Length 165;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ELQVIGLQPTP 12
||:|||||
Db 36 ELQVIGLQPTP 47

RESULT 4
PCT-US01-03080-8
Sequence 8, Application PC/TUS0103080

GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Bcl-2-like Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: P0044PCT
CURRENT APPLICATION NUMBER: PCT/US01/03080
CURRENT FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/179,487
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/180,697
PRIOR FILING DATE: 2000-02-07
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 195
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-03080-8

Query Match 71.7%; Score 43; DB 1; Length 195;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ELQVIGLQPTP 12
||:|||||
Db 36 ELQVIGLQPTP 47

RESULT 5
US-60-258-251-43
Sequence 43, Application US/60258251
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
AND USES THEREOF
FILE REFERENCE: C0001039
CURRENT APPLICATION NUMBER: US/60/258,251
CURRENT FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 102
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 214
TYPE: PRT
ORGANISM: HUMAN
NAME/KEY: VARIANT
LOCATION: (1)..
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-258-251-43

Query Match 65.0%; Score 39; DB 23; Length 214;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ELQVIGLQPTP 12
||:|||||
Db 48 ELQVIGLQPTP 59

RESULT 6
US-09-467-008-6
Sequence 6, Application US/09467008
GENERAL INFORMATION:
APPLICANT: TRIEBEL, Frederic
MASTRANGELI, Renato
ROMAGNANI, Sergio
TITLE OF INVENTION: LAG-3 SPLICED VARIANTS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESS: BROWDY AND NEIMARK, P.L.L.C.
STREET: 624 Ninth Street, N.W., Suite 300

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/467,008
FILING DATE: 20-Dec-1999
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 97401404.5
FILING DATE: 18-JUN-1997
APPLICATION NUMBER: PCT/EP98/03307
FILING DATE: 03-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: TRIEBEL-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 338 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-467-008-6

Query Match 65.0%: Score 39; DB 18; Length 338;
Best Local Similarity 72.7%: Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LQVIGLQLEPTP 12
| | | | |
DB 235 LTVIGLEPTP 245

RESULT 7
US-07-922-466-6
Sequence 6, Application US/07922466
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
TITLE OF INVENTION: New Proteins Produced By Human
TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/922,466
FILING DATE: 19920710
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0725.0020001/JAG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)466-0800
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-922-466-6

Query Match 65.0%: Score 39; DB 3; Length 380;
Best Local Similarity 72.7%: Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LQVIGLQLEPTP 12
| | | | |
DB 257 LTVIGLEPTP 267

RESULT 8
US-07-922-466A-6
Sequence 6, Application US/07922466A
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
TITLE OF INVENTION: New Proteins Produced By Human
TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/922,466A
FILING DATE: 10-JUL-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0725.0020001/JAG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-922-466A-6

Query Match 65.0%: Score 39; DB 3; Length 380;
Best Local Similarity 72.7%: Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LQVIGLQLEPTP 12
| | | | |
DB 257 LTVIGLEPTP 267

TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-988-6

Query Match 65.0%; Score 39; DB 8; Length 380;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LQVIGLQLEPTP 12
Db 257 LTVIGLEPPPT 267

RESULT 12
US-08-474-988A-6
Sequence 6, Application US/08474988A.
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,988A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/416,478
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-988A-6

Query Match 65.0%; Score 39; DB 8; Length 380;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 2 LQVIGLQLEPTP 12

Db 257 LTVIGLEPPPT 267

RESULT 13
US-09-467-008-4
Sequence 4, Application US/09467008
GENERAL INFORMATION:
APPLICANT: TRIEBEL, Frederic
APPLICANT: MASTRANGELI, Renato
APPLICANT: ROMAGNANI, Sergio
TITLE OF INVENTION: LAG-3 SPLICING VARIANTS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/467,008
FILING DATE: 20-Dec-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 97401404.5
FILING DATE: 18-JUN-1997
APPLICATION NUMBER: PCT/EP98/03307
FILING DATE: 03-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: TRIEBEL-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-467-008-4

Query Match 65.0%; Score 39; DB 18; Length 422;
Best Local Similarity 72.7%; Pred. No. 2.4e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LQVIGLQLEPTP 12
Db 235 LTVIGLEPPPT 245

RESULT 14
US-08-394-442-7
Sequence 7, Application US/08394442
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,442
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: HERCEND-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 470 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-394-442-7

Query Match 65.0%; Score 39; DB 7; Length 470;
Best Local Similarity 72.7%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 LOYVIGLPP 12
| | | | | | | | | |
Db 229 LTVGLPP 239

RESULT 15
US-08-474-988-7
Sequence 7 Application US/08474988
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,988
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/416,478
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:

NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 470 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-988-7

Query Match 65.0%; Score 39; DB 8; Length 470;
Best Local Similarity 72.7%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 LOYVIGLPP 12
| | | | | | | | | |
Db 229 LTVGLPP 239

Search completed: June 4, 2001, 12:20:05
Job time: 518 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:21:29 ; Search time 84.28 Seconds

(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-28

Perfect score: 60

Sequence: 1 ELQVGLQLPTP 12

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA.New:*

```
1: /cgn2_6/ptodata/1/paa/PCr_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	12	5	US-09-284-107-28
2	38	63.3	141	1	PCT-US01-01312-1035
3	35	58.3	110	1	PCT-US01-01339-2813
4	35	58.3	110	1	PCT-US01-01329-1005
5	34	56.7	89	1	PCT-US01-11988-2091
6	34	56.7	89	5	US-09-833-245-2091
7	34	56.7	102	1	PCT-US01-01339-3249
8	34	56.7	127	1	PCT-US01-11988-2090
9	34	56.7	127	5	US-09-833-245-2090
10	34	56.7	335	5	US-09-739-449-9638
11	34	56.7	350	1	PCT-US01-04098A-1405
12	34	56.7	843	5	US-09-739-449-13150
13	33	55.0	23	5	US-09-809-391-328
14	33	55.0	257	5	US-09-178-115-51
15	33	55.0	334	1	PCT-US01-01332-1196
16	33	55.0	377	1	US-09-178-115-87
17	33	55.0	459	5	US-09-178-115-2
18	33	55.0	548	1	PCT-US01-01332-919
19	32	53.3	51	5	US-09-828-769-283
20	32	53.3	190	5	US-09-270-849B-180083
21	32	53.3	259	5	US-09-739-449-12942
22	32	53.3	303	5	US-09-739-449-8169
23	32	53.3	347	1	PCT-US01-11988-1273
24	32	53.3	347	1	PCT-US01-11988-1274
25	32	53.3	347	1	PCT-US01-11988-1275
26	32	53.3	347	5	US-09-833-245-1273
27	32	53.3	347	5	US-09-833-245-1274

28	32	53.3	347	5	US-09-833-245-1275	Sequence 1275, Ap
29	32	53.3	425	5	US-09-739-449-8808	Sequence 8808, Ap
30	32	53.3	499	6	US-60-248-505-882	Sequence 882, App
31	32	53.3	546	5	US-09-739-449-12091	Sequence 12091, A
32	32	53.3	555	5	US-09-813-872-4	Sequence 4, Appl1
33	32	53.3	571	5	US-09-739-449-9005	Sequence 9005, Ap
34	32	53.3	582	1	PCT-US01-04098A-1708	Sequence 1708, Ap
35	32	53.3	617	1	PCT-US01-04098A-3676	Sequence 3676, Ap
36	32	53.3	714	6	US-60-248-505-1225	Sequence 1225, Ap
37	32	53.3	890	5	US-09-270-849B-181560	Sequence 181560, Ap
38	32	53.3	891	6	US-60-248-505-797	Sequence 797, App
39	32	53.3	972	5	US-09-628-359-17	Sequence 17, Appl
40	32	53.3	1294	6	US-60-248-505-1177	Sequence 1177, Ap
41	31	51.7	49	5	US-09-814-666-369	Sequence 369, App
42	31	51.7	99	1	PCT-US01-11988-893	Sequence 893, App
43	31	51.7	99	5	US-09-833-245-893	Sequence 893, App
44	31	51.7	113	5	US-09-270-849B-184149	Sequence 184149, Ap
45	31	51.7	140	1	PCT-US01-01339-4275	Sequence 4275, Ap

ALIGNMENTS

```
RESULT 1
US-09-284-107-28
Sequence 28, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: AND PROTEINS HAVING SPECIFIC AFFINITY FOR A TARGET
FILE REFERENCE: 31632000600
CURRENT APPLICATION NUMBER: US/09/284,107
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
PRIOR FILING DATE: 1996-10-08
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligopeptides representing the extracellular
US-09-284-107-28
Query Match
Best Local Similarity 100.0%; Score 60; DB 5; Length 12;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ELQVGLQLPTP 12
DB 1 ELQVGLQLPTP 12
RESULT 2
PCT-US01-01312-1035
Sequence 1035, Application PC/TUS0101312
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT332PCT
CURRENT APPLICATION NUMBER: PCT/US01/01312
CURRENT FILING DATE: 2001-05-09
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1035
LENGTH: 141
```

```
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (124)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01312-1035
```

```
Query Match
Best Local Similarity 63.3%; Score 38; DB 1; Length 141;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 2 LQVLGLQLPTP 12
    || || || |
Db 74 LQVLGLSLPLP 84
```

```
RESULT 3
PCT-US01-01339-2813
Sequence 2813, Application PC/TUS0101339
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006PCT
CURRENT APPLICATION NUMBER: PCT/US01/01339
CURRENT FILING DATE: 2001-03-17
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2813
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids.
PCT-US01-01339-2813
```

```
Query Match
Best Local Similarity 58.3%; Score 35; DB 1; Length 110;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 3 QVLGLQLPTP 12
    || | | | |
Db 49 QVPGFQPTPT 58
```

```
RESULT 4
PCT-US01-01329-1005
Sequence 1005, Application PC/TUS0101329
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA120PCT
CURRENT APPLICATION NUMBER: PCT/US01/01329
CURRENT FILING DATE: 2001-01-17
NUMBER OF SEQ ID NOS: 3506
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1005
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01329-1005
```

```
Query Match
Best Local Similarity 58.3%; Score 35; DB 1; Length 110;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 3 QVLGLQLPTP 12
    || | | | |
Db 49 QVPGFQPTPT 58
```

```
RESULT 5
PCT-US01-11988-2091
Sequence 2091, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2091
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (87)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-11988-2091
```

```
Query Match
Best Local Similarity 56.7%; Score 34; DB 1; Length 89;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 ELQVVLGLQLP 10
    || | | | | |
Db 30 ELKVLGLSLP 39
```

```
RESULT 6
US-09-833-245-2091
Sequence 2091, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2091
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (87)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-2091
```

```
Query Match
Best Local Similarity 56.7%; Score 34; DB 1; Length 89;
```

Best Local Similarity 60.0%; Pred. No. 79;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 10
||:||||:
Db 30 ELEKLGLEIP 39

RESULT 7
PCT-US01-01339-3249
; Sequence 3249, Application PC/TUS0101339
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01339
; CURRENT FILING DATE: 2001-03-17
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3249
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01339-3249

Query Match 56.7%; Score 34; DB 1; Length 102;
Best Local Similarity 58.3%; Pred. No. 90;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 12
|:|:|:|:
Db 11 EKVVGRLPTP 22

RESULT 8
PCT-US01-11988-2090
; Sequence 2090, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (109)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (116)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-11988-2090

Query Match 56.7%; Score 34; DB 1; Length 127;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 10
||:||||:
Db 30 ELEKLGLEIP 39

RESULT 9
US-09-833-245-2090
; Sequence 2090, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (109)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (116)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-2090

Query Match 56.7%; Score 34; DB 5; Length 127;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 10
||:||||:
Db 30 ELEKLGLEIP 39

RESULT 10
US-09-739-449-9638
; Sequence 9638, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15480)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9638
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9638

Query Match 56.7%; Score 34; DB 5; Length 335;
Best Local Similarity 58.3%; Pred. No. 3.1e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 12

Db 246 ELVDVMDQTLPTP 257

RESULT 11

PCT-US01-04098A-1405
; Sequence 1405, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-028
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1405
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1405

Query Match 56.7%; Score 34; DB 1; Length 350;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ELQVGLQPLPTP 12
:|:|:|:|:|:|
Db 220 KLEVYDLDQISRP 231

RESULT 12

US-09-739-449-13150
; Sequence 13150, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 13150
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-13150

Query Match 56.7%; Score 34; DB 5; Length 843;
Best Local Similarity 55.6%; Pred. No. 7.9e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLGLQPLPTP 12
|:|:|:|:|:|
Db 425 VIGIELPPTP 433

RESULT 13

US-09-809-391-328
; Sequence 328, Application US/09809391
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 328
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-391-328

Query Match 55.0%; Score 33; DB 5; Length 23;
Best Local Similarity 77.8%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VLGLQPLPTP 12
|:|:|:|:|:|
Db 9 VLGLLPTP 17

RESULT 14

US-09-178-115-51
; Sequence 51, Application US/09178115
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/178,115
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 09/177,776
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PY-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51

LENGTH: 257
TYPE: PRT
ORGANISM: HUMAN
US-09-178-115-51

Query Match 55.0%; Score 33; DB 5; Length 257;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LOVIGLQLP 10
|:|:|:|:|:
Db 46 LELIGFQLP 54

RESULT 15
PCT-US01-01332-1196
Sequence 1196, Application PC/TUS0101332
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P0202PCT
CURRENT APPLICATION NUMBER: PCT/US01/01332
CURRENT FILING DATE: 2001-05-09
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1249
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1196
LENGTH: 334
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (184)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01332-1196

Query Match 55.0%; Score 33; DB 1; Length 334;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 VLGLOLPTP 12
|:|:|:|:|:
Db 45 VKGINLPTP 53

Search completed: June 4, 2001, 12:21:30
Job time: 592 sec


```

PCN-US00-05882-1128
; Sequence 1128, Application PC/TUS00058832
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben

```

1 TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
 2
 3 FILE REFERENCE: PA106PCP
 4
 5 CURRENT APPLICATION NUMBER: PCT/US00/05882
 6
 7 CURRENT FILING DATE: 2000-03-08
 8
 9 EARLIER APPLICATION NUMBER: 60/124,270
 10
 11 EARLIER FILING DATE: 1999-03-12
 12
 13 NUMBER OF SEQ ID NOS: 1694
 14
 15 SOFTWARE: PatentIn Ver. 2.0
 16
 17 SEQ ID NO 1128
 18
 19 LENGTH: 399
 20
 21 TYPE: PRT
 22
 23 ORGANISM: Homo sapiens
 24
 25 FEATURE:
 26
 27 NAME/KEY: SITE
 28
 29 LOCATION: (208)
 30
 31 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 32
 33 FEATURE:
 34
 35 NAME/KEY: SITE
 36
 37 LOCATION: (349)
 38
 39 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 40
 41 PCT-US00-05882-1128

Query Match	100.0%	Score 68;	DB 1;	Length 399;
Best Local Similarity	100.0%;	Pred. No. 0.0088;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	VWEHYLFYLAAG	12	
			5	
Db	314	VWEHYLFYLAAG	325	

RESULT 3
US-07-661-95-25
Sequence 25, Application US/07661895
GENERAL INFORMATION:
APPLICANT: HOGARTH, Phillip M.
APPLICANT: HULETT, Mark D.
APPLICANT: IERINO, Francesco L.
APPLICANT: MCKENZIE, Ian F. C.
APPLICANT: OSMAN, Nalin
TITLE OF INVENTION: HYBRID FC RECEPTOR MOLECULES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
City: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/861,895
FILING DATE: 19920624
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00513
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7045
FILING DATE: 25-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17227/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 25:

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 410 amino acids
;     TYPE: AMINO ACID
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;
US-07-861-895-25

```

Query Match	94.1%;	Score 64;	DB 3;	Length 410;
Best Local Similarity	83.3%;	Pred: No. 0.037;		
Matches 10; Conservative	2;	Mismatches	0;	Indels 0; Gaps 0;

```
QY      1 VWFHVFYLVG 12
          ||||:||||:|
Db      304 VWFHILFYLSVG 315
```

```

RESULT      4
US-09-252-691-7770
: Sequence 7770, Application US/09252691B
:
: GENERAL INFORMATION:
: APPLICANT: Keith G. Welnstock et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTE
: TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS.
: FILE REFERENCE: 107196.135
: CURRENT APPLICATION NUMBER: US/09/252.691B
: CURRENT FILING DATE: 1999-02-18
: NUMBER OF SEQ ID NOS: 11324
: SEQ ID NO 7770
:
: LENGTH: 370
:
: TYPE: PR1
:
: ORGANISM: Enterobacter cloacae
US-09-252-691-7770

```

Query Match	61.88;	Score 42;	DB 16;	Length 370;
Best Local Similarity	66.78;	Pred. No. 91;		
Matches 6;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      2 WEHVLEFLA 10
        1 : 1 : 1 1 1
Db      131 WYHGIFFLA 139
```

```

1 RESULT 5
2 US-09-252-691C-7770
3 : Sequence 7770, Application US/09252691C
4 : GENERAL INFORMATION:
5 : APPLICANT: Keith G. Weinstein et al.
6 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
7 : -TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
8 : FILE REFERENCE: 107196.135
9 : CURRENT APPLICATION NUMBER: US/09/252,691C
10 : PRIOR FILING DATE: 1999-02-18
11 : PRIOR APPLICATION NUMBER: US 60/094,145
12 : PRIOR FILING DATE: 1998-07-24
13 : PRIOR APPLICATION NUMBER: US 60/074,787
14 : PRIOR FILING DATE: 1998-02-18
15 : NUMBER OF SEQ. ID NOS: 11326
16 : SEQ. ID NO 7770.
17 : LENGTH: 370
18 : TYPE: PRT
19 : ORGANISM: Enterobacter cloacae
20 : US-09-252-691C-7770

```

Query Match	61.88;	Score 42;	DB 16;	Length 370;
Best Local Similarity	66.78;	Pred. No. 91;		
Matches	6;	Conservative	2;	Mismatches 1;
				Indels 0;
				Gaps 0;
OY	2	WFHVFLEYLA	10	
		: :		
db	131	WYHGIFLYLA	139	

```

: RESULT 6
: US-09-489-039A-7598
: Sequence 7598, Application US/09489039A
: GENERAL INFORMATION:
: APPLICANT: Gary Breton et. al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
: TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.2004001
: CURRENT APPLICATION NUMBER: US/09/489,039A
: CURRENT FILING DATE: 2000-01-27
: PRIOR APPLICATION NUMBER: US 60/117,747
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 14342
: SEQ ID NO 7598
: LENGTH: 410
: TYPE: prt
: ORGANISM: Klebsiella pneumoniae
: US-09-489-039A-7598

```

Query Match	61.8%	Score 42	DB 18	Length 410
Best Local Similarity	45.5%	Pred. No. 99		
Matches	5	Conservative	4	Mismatches
			2	Indels
				Gaps
QY	2 WFWHVLFLANG	12		
	: : :			
DB	160 WFWNLVYLQIG	170		

RESULT 7
 US-09-107-433-2896
 Sequence 2896, Application US/09107433
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
 SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
 THERAPEUTICS
 NUMBER OF SEQUENCES: 5206
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: <unknown>
 OPERATING SYSTEM: <unknown>
 SOFTWARE: <unknown>
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,433
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085131
 FILING DATE: May 12, 1998
 APPLICATION NUMBER: 60/051553
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariofello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-011
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 2896:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 104 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

```

?      :
?      :   HYPOTHETICAL: YES
?      :   ORIGINAL SOURCE:
?      :   ORGANISM: Streptococcus pneumoniae
?      :
?      :   FEATURE:
?      :       NAME/KEY: misc_feature
?      :       LOCATION: (B) LOCATION 1...104
?      :
?      :   SEQUENCE DESCRIPTION: SEQ ID NO: 2896:
US-09-107-433-2896

```

Query Match	60.3%	Score 41;	DB 15;	Length 104;
Best Local Similarity	60.0%	Pred. No. 49;		
Matches	6;	Conservative	2;	Mismatches 2;
				Indels 0;
				Gaps 0;
Oy	2	WERVLEFLVAV	11	
		:	11:	
Db	3	WERHLEFLVAV	12	

```

RESULT      8
US-09-450-969-6248
; Sequence 6248, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH99-09A
; CURRENT APPLICATION NUMBER: US/09/450.969
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 6248
; LENGTH: 244
; TYPE: prt
; ORGANISM: S.epidermidis
US-09-450-969-6248

```

	Query Match	Similarity	60.3%	Score 41	DB 18	Length 244
	Best Local	Similarity	41.7%	Pred. No.	94	
Matches	5	Conservative	5	Mismatches	2	Gaps 0
QY	1	VWFHVFYIANG	12			
		:	1			
DB	139	IMEYLFIFLEFG	150			

```

RESULT      9
US-09-252-691-11281
; Sequence 11281, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691B
; CURRENT FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 11281
; LENGTH: 418
; TYPE: prt
; ORGANISM: Enterobacter cloacae
US-09-252-691-11281

```

```

Query Match Similarity 60.3%; Score 41; DB 16; Length 418;
Best Local Similarity 45.5%; Pred No. 1.4e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0.
QY 2 WPHVLEFLAVG 12
||::: ||: |
Db 168 WENIVLYLDIG 178

```

```
RESULT 10
US-09-252-691C-11281
; Sequence 11281, Application US/09252691C
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691C
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/094,145
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: US 60/074,787
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 11326
; SEQ ID NO 11281
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-09-252-691C-11281

Query Match
Best local Similarity 60.3%; Score 41; DB 16; Length 418;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 WFWHLYFLANG 12
DB 168 WFWHLYFLANG 178

RESULT 11
US-09-417-507-27932
; Sequence 27932, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PAT99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 27932
; LENGTH: 497
; TYPE: PRT
; ORGANISM: A. fumigatus
US-09-417-507-27932

Query Match
Best local Similarity 60.3%; Score 41; DB 18; Length 497;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 FHWLYFLANG 12
DB 27 FHWLYFLANG 36

RESULT 12
US-09-270-767-35611
; Sequence 35611, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 35611
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
```

```
FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35611

Query Match
Best local Similarity 58.8%; Score 40; DB 16; Length 239;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 WFWHLYFL 9
DB 224 WFWHLYFL 231

RESULT 13
US-09-270-767-50828
; Sequence 50828, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 50828
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50828
```

```
Query Match
Best local Similarity 58.8%; Score 40; DB 16; Length 239;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 WFWHLYFL 9
DB 224 WFWHLYFL 231

RESULT 14
US-09-417-507-28756
; Sequence 28756, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PAT99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 28756
; LENGTH: 80
; TYPE: PRT
; ORGANISM: A. fumigatus
US-09-417-507-28756

Query Match
Best local Similarity 57.4%; Score 39; DB 18; Length 80;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WFWHLYFLANG 12
DB 24 WFWHLYFLANG 35

RESULT 15
US-09-741-669-400
; Sequence 400, Application US/09741669
```

GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allyn
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
TITLE OF INVENTION: Genes identified as required for
FILE REFERENCE: ELITRA 009A
CURRENT APPLICATION NUMBER: US/09/741,669
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 400
LENGTH: 421
TYPE: PRT
ORGANISM: Escherichia coli
US-09-741-669-400

Query Match 57.4%, Score 39, DB 21, Length 421;
Best Local Similarity 50.0%; Pred. No. 3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VMEHVLFYIANG 12
Db 171 IMFYSLVYIANG 182

Search completed: June 4, 2001, 12:20:06
Job time: 519 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:21:30 ; Search time 84.28 Seconds
(without alignments)
2.148 Million cell updates/sec

Title: US-09-284-107-29

Perfect score: 68

Sequence: 1 VWFHLYFLAVG 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 296725 seqs, 15086421 residues

Total number of hits satisfying chosen parameters: 296725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	12	5	US-09-284-107-29
2	41	60.3	36	5	US-09-739-449-9245
3	40	58.8	239	5	US-09-270-8498-183288
4	40	58.8	331	5	US-09-739-449-11606
5	38	55.9	365	5	US-09-270-8498-185951
6	37	54.4	61	5	US-09-270-8498-193851
7	37	54.4	141	5	US-09-270-8498-182343
8	36	52.9	870	5	US-09-270-8498-189447
9	36	52.9	1488	6	US-60-248-505-811
10	36	52.9	1535	6	US-60-248-505-1121
11	35	51.5	53	1	PCT-US01-01339-5081
12	35	51.5	77	1	PCT-US01-01335-353
13	35	51.5	95	1	PCT-US01-04098A-1850
14	35	51.5	123	5	US-09-270-8498-187144
15	34	50.0	62	5	US-09-270-8498-186130
16	34	50.0	101	5	US-09-270-8498-189207
17	34	50.0	124	5	US-09-811-284-165
18	34	50.0	159	5	US-09-811-284-168
19	34	50.0	195	5	US-09-811-284-241
20	34	50.0	1879	1	PCT-US00-35017A-1265
21	34	50.0	1960	1	PCT-US01-04098A-1516
22	34	50.0	1963	1	PCT-US01-04098A-1854
23	33.5	49.3	258	1	PCT-US01-04098A-1952
24	33	48.5	37	1	PCT-US01-11988-662
25	33	48.5	37	1	PCT-US01-11988-663
26	33	48.5	37	5	US-09-833-245-662
27	33	48.5	37	5	US-09-833-245-663

28	33	48.5	46	1	PCT-US01-01321-977	Sequence 977, App
29	33	48.5	66	5	US-09-739-449-9314	Sequence 9314, App
30	33	48.5	92	1	US-09-814-666-364	Sequence 364, App
31	33	48.5	191	5	PCT-US01-04098A-3818	Sequence 3818, App
32	33	48.5	227	5	US-09-270-8498-184164	Sequence 184164, App
33	33	48.5	266	5	US-09-270-8498-192136	Sequence 192136, App
34	33	48.5	273	1	PCT-US01-01312-968	Sequence 968, App
35	33	48.5	331	1	PCT-US01-11988-1990	Sequence 1990, App
36	33	48.5	331	5	US-09-833-245-1990	Sequence 1990, App
37	33	48.5	407	5	US-09-739-449-8921	Sequence 8921, App
38	33	48.5	467	5	PCT-US01-04098A-1004	Sequence 1004, App
39	33	48.5	467	5	US-09-423-844-195	Sequence 195, App
40	33	48.5	467	5	US-09-380-139A-195	Sequence 195, App
41	33	48.5	496	1	PCT-US01-04098A-2972	Sequence 2972, App
42	33	48.5	518	5	US-09-739-449-13267	Sequence 13267, App
43	33	48.5	596	1	PCT-US01-04098A-1605	Sequence 1605, App
44	33	48.5	607	1	PCT-US01-04098A-3573	Sequence 3573, App
45	33	48.5	635	5	US-09-739-449-9951	Sequence 9951, App

ALIGNMENTS

```
RESULT 1
US-09-284-107-29
Sequence 29, Application US/09284107
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
TITLE OF INVENTION: METHODS AND MEANS FOR SELECTING PEPTIDES
TITLE OF INVENTION: AND PROTEINS HAVING SPECIFIC AFFINITY FOR A TARGET
FILE REFERENCE: 313632000600
CURRENT APPLICATION NUMBER: US/09/284,107
PRIOR FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: PCT/NL97/00557
PRIOR FILING DATE: 1997-10-07
PRIOR APPLICATION NUMBER: EP 96202791.8
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: domain of CD64
US-09-284-107-29
Query Match 100.0%; Score 68; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 VWFHLYFLAVG 12
Db 1 VWFHLYFLAVG 12
RESULT 2
US-09-739-449-9245
Sequence 9245, Application US/09739449
GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10/115490/C
CURRENT FILING DATE: 2000-12-19
CURRENT APPLICATION NUMBER: US/09/739,449
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 9245
```

LENGTH: 36
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9245

Query Match 60.3%; Score 41; DB 5; Length 36;
Best Local Similarity 63.6%; Pred. No. 5.2;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 WPHVLFYLVG 12
|||
Db 4 WFRALFYLPFG 14

RESULT 3
US-09-270-849B-183288
Sequence 183288, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 183288
LENGTH: 239
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-183288

Query Match 58.8%; Score 40; DB 5; Length 239;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 WPHVLFYLV 9
|||
Db 224 WFRALFYLV 231

RESULT 4
US-09-739-449-11606
Sequence 11606, Application US/09739449
GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15490)C
CURRENT APPLICATION NUMBER: US/09/739,449
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 11606
LENGTH: 331
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11606

Query Match 58.8%; Score 40; DB 5; Length 331;
Best Local Similarity 54.5%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 WPHVLFYLV 11
|||
Db 202 WPHVLFYLV 212

RESULT 5
US-09-270-849B-185951
Sequence 185951, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 185951
LENGTH: 365
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-185951

Query Match 55.9%; Score 38; DB 5; Length 365;
Best Local Similarity 50.0%; Pred. No. 93;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 WPHVLFYLVG 12
|||
Db 123 ITHILYLVG 134

RESULT 6
US-09-270-849B-193851
Sequence 193851, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 193851
LENGTH: 61
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-193851

Query Match 54.4%; Score 37; DB 5; Length 61;
Best Local Similarity 54.5%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPHVLFYLV 11
|||
Db 5 IFFVLYLV 15

RESULT 7
US-09-270-849B-182343
Sequence 182343, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 182343
LENGTH: 141

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid.
US-09-270-849B-182343

Query Match 54.4%; Score 37; DB 5; Length 141;
Best Local Similarity 57.1%; Pred. No. 65;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 WPHVLEFY 8
DB 105 WPHVLEFY 111

RESULT 8
US-09-270-849B-189447
Sequence 189447, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 189447
LENGTH: 870
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-189447

Query Match 52.9%; Score 36; DB 5; Length 870;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WPHVLEFY 7
DB 744 WPHVLEFY 750

RESULT 9
US-60-248-505-811
Sequence 811, Application US/60248505
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: C1000918
CURRENT APPLICATION NUMBER: US/60/248,505
CURRENT FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 1998
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 811
LENGTH: 1488
TYPE: PRT
ORGANISM: Human
US-60-248-505-811

Query Match 52.9%; Score 36; DB 6; Length 1488;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 WPHVLEFY 9
DB 111 WPHVLEFY 111

DB 1266 WPHVLEFY 1273

RESULT 10
US-60-248-505-1121
Sequence 1121, Application US/60248505
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: C1000918
CURRENT APPLICATION NUMBER: US/60/248,505
CURRENT FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 1998
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1121
LENGTH: 1535
TYPE: PRT
ORGANISM: Human
US-60-248-505-1121

Query Match 52.9%; Score 36; DB 6; Length 1535;
Best Local Similarity 50.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 WPHVLEFY 9
DB 1255 WPHVLEFY 1262

RESULT 11
PCT-US01-01339-5081
Sequence 5081, Application PC/TUS0101339
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006PCT
CURRENT APPLICATION NUMBER: PCT/US01/01339
CURRENT FILING DATE: 2001-03-17
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5081
LENGTH: 53
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-01339-5081

Query Match 51.5%; Score 35; DB 1; Length 53;
Best Local Similarity 40.0%; Pred. No. 65;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 WPHVLEFY 10
DB 24 WPHVLEFY 33

RESULT 12
PCT-US01-01335-353
Sequence 353, Application PC/TUS0101335
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC001PCT
CURRENT APPLICATION NUMBER: PCT/US01/01335
CURRENT FILING DATE: 2001-01-17
NUMBER OF SEQ ID NOS: 879
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 353
LENGTH: 77
TYPE: PRT

OY 2 WPHVLEFY 9
DB 111 WPHVLEFY 111

ORGANISM: Homo sapiens
PCT-US01-01335-353

Query Match 51.5%; Score 35; DB 1; Length 77;
Best Local Similarity 42.9%; Pred. No. 86;
Matches 6; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

OY 1 VWF--HVLFLAVG 12
Db 47 IWLPHLFLAVG 60

RESULT 13

PCT-US01-04098A-1850
Sequence 1850, Application PC/TUS0104098A
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: PCT/US01/04098A
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 1850
LENGTH: 95
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-04098A-1850

Query Match 51.5%; Score 35; DB 1; Length 95;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 HVLFLAVG 12
Db 37 HVLFLAVG 45

RESULT 14

US-09-270-849B-187144
Sequence 187144, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 187144
LENGTH: 123
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-187144

Query Match 51.5%; Score 35; DB 5; Length 123;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VWFHVLFL 9
Db 101 VWFHVLFL 109

RESULT 15

US-09-270-849B-186130
Sequence 186130, Application US/09270849B
GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270,849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 186130
LENGTH: 62
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-186130

Query Match 50.0%; Score 34; DB 5; Length 62;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VWFHVLFL 8
Db 28 LMVYVLFY 35

Search completed: June 4, 2001, 12:21:30
Job time: 592 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:10:07 ; Search time 106.14 Seconds
(without alignments)
6.463 Million cell updates/sec

Title: US-09-284-107-21

Perfect score: 55
Sequence: 1 VLNAVSTSPLE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.0401:*

- 1: /SID56/gcgdata/geneseq/geneseqp/AA1980.DAT:*
- 2: /SID56/gcgdata/geneseq/geneseqp/AA1981.DAT:*
- 3: /SID56/gcgdata/geneseq/geneseqp/AA1982.DAT:*
- 4: /SID56/gcgdata/geneseq/geneseqp/AA1983.DAT:*
- 5: /SID56/gcgdata/geneseq/geneseqp/AA1984.DAT:*
- 6: /SID56/gcgdata/geneseq/geneseqp/AA1985.DAT:*
- 7: /SID56/gcgdata/geneseq/geneseqp/AA1986.DAT:*
- 8: /SID56/gcgdata/geneseq/geneseqp/AA1987.DAT:*
- 9: /SID56/gcgdata/geneseq/geneseqp/AA1988.DAT:*
- 10: /SID56/gcgdata/geneseq/geneseqp/AA1989.DAT:*
- 11: /SID56/gcgdata/geneseq/geneseqp/AA1990.DAT:*
- 12: /SID56/gcgdata/geneseq/geneseqp/AA1991.DAT:*
- 13: /SID56/gcgdata/geneseq/geneseqp/AA1992.DAT:*
- 14: /SID56/gcgdata/geneseq/geneseqp/AA1993.DAT:*
- 15: /SID56/gcgdata/geneseq/geneseqp/AA1994.DAT:*
- 16: /SID56/gcgdata/geneseq/geneseqp/AA1995.DAT:*
- 17: /SID56/gcgdata/geneseq/geneseqp/AA1996.DAT:*
- 18: /SID56/gcgdata/geneseq/geneseqp/AA1997.DAT:*
- 19: /SID56/gcgdata/geneseq/geneseqp/AA1998.DAT:*
- 20: /SID56/gcgdata/geneseq/geneseqp/AA1999.DAT:*
- 21: /SID56/gcgdata/geneseq/geneseqp/AA2000.DAT:*
- 22: /SID56/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	12	19 W60553	Oligopeptide from
2	55	100.0	261	20 Y33183	Human sfc-gammari
3	55	100.0	344	13 R22549	Human macrophage-s
4	55	100.0	344	17 R91439	Human FCRI (CDNA C
5	55	100.0	344	21 Y66183	Human macrophage-s
6	55	100.0	374	13 R20811	Human macrophage-s
7	55	100.0	374	13 R22550	Human macrophage-s
8	55	100.0	374	17 R91438	Human FCRI (CDNA C
9	55	100.0	374	17 W00859	Human FCRI (CDNA C
10	55	100.0	374	19 W80448	Human FC receptor
11	55	100.0	374	19 W97833	Human FC receptor

12	55	100.0	374	19 W97834	Human FC receptor
13	55	100.0	374	21 Y66134	Human macrophage-s
14	55	100.0	374	21 Y96226	Human high affinity
15	55	100.0	399	21 B43683	Human cancer assoc
16	36	65.5	292	21 B28636	Human breast tumor
17	36	65.5	329	21 B28626	Human B11a1 anti
18	36	65.5	329	21 Y82015	Human immunogenic
19	34	61.8	126	22 B63956	Human prostate can
20	34	61.8	187	19 W34052	C. parvum p23 prot
21	34	61.8	322	20 W85718	Novel protein (C10
22	34	61.8	364	21 G47084	Arabidopsis thalia
23	34	61.8	365	21 G47111	Arabidopsis thalia
24	34	61.8	374	21 G47083	Arabidopsis thalia
25	34	61.8	375	21 G47110	Arabidopsis thalia
26	34	61.8	410	12 R12428	Hybrid Fc(gamma)R
27	34	61.8	442	21 G47082	Arabidopsis thalia
28	34	61.8	443	21 G47109	Arabidopsis thalia
29	33	60.0	94	21 G35306	Zea mays protein f
30	33	60.0	131	19 W58851	Human AC222.1 prot
31	33	60.0	131	19 W37144	Human AC222.1 prot
32	33	60.0	139	21 G35305	Zea mays protein f
33	33	60.0	144	21 G40794	Zea mays protein f
34	33	60.0	350	8 P70645	Pseudotables: virus
35	33	60.0	350	15 R63143	Glycoprotein 63 (g
36	33	60.0	377	20 W86195	Human FC receptor
37	33	60.0	730	9 P80618	Human Bone Morphog
38	33	60.0	780	18 W13669	C-proteinase encod
39	33	60.0	788	19 W75919	C-proteinase sequ
40	33	60.0	986	18 W13670	C-proteinase encod
41	32	58.2	139	21 G04555	Arabidopsis thalia
42	32	58.2	139	21 G37898	Arabidopsis thalia
43	32	58.2	176	21 G20314	Arabidopsis thalia
44	32	58.2	176	21 G38740	Arabidopsis thalia
45	32	58.2	181	21 G20313	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
ID W60553	W60553 standard; peptide: 12 AA.
XX W60553:	
XX 18-AUG-1998 (first entry)	
XX Oligopeptide from extracellular domain of CD64.	
XX DE	
XX KW Extracellular domain; CD64; identification: antibody;	
XX KW Immunohistochemical; immunofluorescent analysis; detection:	
XX KW cell transformation; mutation; anti; oncogene.	
XX OS	
XX OS Synthetic.	
XX PN W09815833-A1.	
XX PD 16-APR-1998.	
XX PF 07-OCT-1997; 97MO-NL00557.	
XX PR 08-OCT-1996; 96EP-0202791.	
XX (UYUT-) RIKUSUNIV UTRCHT.	
XX PA De Krulff CA, Logtenberg T;	
XX PI WPI; 1998-240964/21.	
XX DR	
XX PT Identifying peptide(s) binding specifically to protein target - by	
XX PT expressing on phage surface and testing for binding to immobilised	
XX PT oligo:peptide derived from the target, useful for, e.g. identifying	
XX PT specific antibodies	

XX Example 1; Page 29; 40pp; English.

CC Synthetic oligopeptides W60537-61 are derived from the extracellular
CC domain of CD64. They were synthesised on 25 polyethylene rods as
CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
CC affinity binding to the CD64-covered rods. The specification describes a
CC method for the identification of a peptide able to bind specifically to a
CC target protein. The method comprises displaying the peptide on the
CC surface of a replicable display package, synthesising oligopeptides
CC derived from the target protein on a solid phase, and testing for binding
CC between the peptide and oligopeptides. The method is used to screen large
CC peptide libraries, especially to detect antibodies, or their fragments,
CC that bind to cell markers or that can differentiate between different
CC forms of the same protein, including bispecific antibodies that bind to
CC two non-overlapping epitopes on the same monomeric antigen or two
CC epitopes on different molecules. The genes/oligonucleotides that encode
CC selected peptides can be isolated and used for recombinant production of
CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
CC immunofluorescent analysis, and also to detect cell transformation caused
CC by mutation in (anti)oncogenes.

SQ Sequence 12 AA;

Query Match 100.0%; Score 55; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. NO. 6.2e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNASVTSPLLE 12
| | | | | | | | | | | |
Db 1 vlnasvtsplle 12

RESULT 2

ID Y33183 standard; Protein; 261 AA.

AC Y33183;

DI 15-NOV-1999 (first entry)

DE Human sFc-gammaRI protein fragment.

KW Fc receptor; Fc-gammaRI; human; FCR; model; three-dimension; 3-D;

KW atomic coordinate; bioactive compound design; computer-assisted;

KW drug design; therapy; inhibitor; Fc-gamma; Fc-epsilon; IgG; IgE;

KW tissue damage; hypersensitivity; inflammatory cell recruitment;

KW inflammatory modulator; Fc-gammaRI; immune function regulation;

KW anti-inflammatory; immunoprotective; sFc-gammaRI.

XX Homo sapiens.

OS Homo sapiens.

PN W09940117-A1.

PD 12-AUG-1999.

PE 04-FEB-1999; 99WO-IB00367.

PR 11-SEP-1998; 98US-0099994.

PA 06-FEB-1998; 98US-0073972.

XX (ILEX-) ILEXUS PTY LTD.

XX Baeli JB, Epa V, Garrett TPJ, Hogarth PM, Matthews BR,

PI Maxwell KR, McCarthy TD, McKenzie IFC, Pietersz GA,

XX Powell MS;

XX MPI; 1999-539978/45.

XX Three-dimensional structures and models of Fc receptors, useful in

XX computer-assisted drug design

PS Claim 11; Page 316-317; 326pp; English.

CC This invention describes a novel model of an Fc receptor (FcR) protein
CC representing a three-dimensional (3-D) structure that substantially
CC conforms to the specified atomic coordinates. Computer model images of
CC the FcR can be used to design bioactive chemical compounds, e.g.
CC oligonucleotides, peptides, peptidomimetics and small organic molecules,
CC by computer-assisted methods of drug design. Therapeutic compositions
CC that inhibit the activity of Fc-gammaR or Fc-epsilonR can be used to
CC reduce IgG- or IgE-mediated tissue damage, e.g. IgG- or IgE-mediated
CC hypersensitivity, recruitment of inflammatory cells or release of
CC inflammatory mediators. The therapeutic compositions can also be used to
CC treat disease in animals. The knowledge of the 3-D structure of Fc-gamma
CC R1a and models of other FcR provides a means for designing and producing
CC compounds that regulate immune function and inflammation in an animal,
CC including humans (i.e. structure based drug design). For example,
CC chemical compounds can be designed to block binding of immunoglobulin to
CC an Fc receptor protein using various computer programs and models. The
CC products of the invention have anti-inflammatory and immunoprotective
CC activity. This sequence represents the human sFc-gammaRI protein
CC fragment described in the method of the invention.

SQ Sequence 261 AA;

Query Match 100.0%; Score 55; DB 20; Length 261;
Best Local Similarity 100.0%; Pred. NO. 0.0036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNASVTSPLLE 12
| | | | | | | | | | | |
Db 174 vlnasvtsplle 185

RESULT 3

ID R22549 standard; Protein; 344 AA.

AC R22549;

DI 21-MAY-1992 (first entry)

DE Human macrophage-specific FcRI receptor encoded by clone p98.

KW Rapid immunoselection cloning technique; cell surface antigen;

KW immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN W09201049-A.

PD 23-JAN-1992.

PE 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GENO-) GEN HOSPITAL CORP.

XX Seed B, Aruffo A, Amlot M;

PI MPI; 1992-056864/07.

PA N-PSDB; Q21179.

XX New CD53 cell surface antigen and DNA encoding it - for

XX immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

XX Example 10; Page 94a; 160pp; English.

XX This amino acid sequence was predicted from the cDNA sequence of

XX cDNA clone p98/X2. It differs from the sequence predicted from

XX clone p135 (see Q21178) at position 25; p135 encodes a Ser residue

XX while clone p98/X2 predicts a Thr residue. At position 58, p90

CC (see Q21180) predicts Val and p98/X2 predicts Leu. Sequences
 CC predicted from all 3 clones show the typical features of a type I
 CC integral membrane protein and include a short hydrophobic signal
 CC sequence, a single 21-residue hydrophobic membrane-spanning domain,
 CC and a short, highly charged cytoplasmic domain. The extracellular
 CC portion contains six potential N-linked glycosylation sites and six
 CC Cys residues distributed among three C2 set Ig-related domains.
 CC A fusion protein of FCRI and a receptor ligand will be helpful to
 CC increase the potency of antibodies in therapy.
 XX
 SQ Sequence 344 AA:

Query Match 100.0%; Score 55; DB 13; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLNAVSTSPLE 12
 |||||
 Db 193 vlnavstspile 204

RESULT 4
 ID R91439 standard; Protein; 344 AA.
 XX R91439;
 XX 30-OCT-1996 (first entry)
 XX
 DE Human FCRI (cDNA clone p98 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FCRI; Fc receptor.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 FT MISC-difference 58 /note="amino acid 58 is Leu in clone p135
 FT translated product"

US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 8805-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI; 1996-200279/20.
 DR N-PSDB; T14718.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX
 XX Example 10; Column 55-56; 79pp; English.

CC The amino acid sequence (R91439) of human FCRI was detd. from a
 CC cDNA clone, p98 (T14718), obt'd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. 2 other cDNA clones, p135 (T14717) and p90 (T14719),
 CC coded for a variants of the FCRI sequence; the C-terminal sequence
 CC of the p98 product is truncated compared with those of the p135

CC and p90 products (see also R91438 and W00859). A fusion protein
 CC of FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.
 XX
 SQ Sequence 344 AA:

Query Match 100.0%; Score 55; DB 17; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLNAVSTSPLE 12
 |||||
 Db 193 vlnavstspile 204

RESULT 5
 ID Y96183 standard; Protein; 344 AA.
 XX Y96183;
 XX 19-DEC-2000 (first entry)
 XX
 DE Human macrophage-specific FCRI.
 XX
 KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
 KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.
 XX
 OS Homo sapiens.

Key Location/Qualifiers

FT MISC-difference 15 /note="encoded by GCG"
 FT MISC-difference 38 /note="encoded by ACC"
 FT MISC-difference 50 /note="encoded by CCG"
 FT MISC-difference 51 /note="encoded by ACC"
 FT MISC-difference 55 /note="encoded by CAC"
 FT MISC-difference 56 /note="encoded by CAC"
 FT MISC-difference 56 /note="encoded by TCC"
 FT MISC-difference 60 /note="encoded by CCC"
 FT MISC-difference 64 /note="encoded by CAG"
 FT MISC-difference 82 /note="encoded by CAA"
 FT MISC-difference 116 /note="encoded by GAA"
 FT MISC-difference 117 /note="encoded by GCA"
 FT MISC-difference 122 /note="encoded by TTC"
 FT MISC-difference 123 /note="encoded by ACC"
 FT MISC-difference 126 /note="encoded by CCG"
 FT MISC-difference 129 /note="encoded by CAA"
 FT MISC-difference 134 /note="encoded by AAT"
 FT MISC-difference 136 /note="encoded by GTT"
 FT MISC-difference 139 /note="encoded by CCA"
 FT MISC-difference 140 /note="encoded by AAT"
 FT MISC-difference 213 /note="encoded by AAT"

FT Misc-difference 216 /note= "encoded by CAA"
 FT /note= "encoded by TTC"
 FT Misc-difference 220 /note= "encoded by CCG"
 FT /note= "encoded by AAT"
 FT Misc-difference 268 /note= "encoded by GCG"
 FT /note= "encoded by GCG"
 FT Misc-difference 305 /note= "encoded by AAC"
 FT /note= "encoded by GCG"
 FT Misc-difference 332 /note= "encoded by GCG"
 FT /note= "encoded by GCG"
 FT Misc-difference 333 /note= "encoded by GCG"
 FT /note= "encoded by GCG"
 FT Misc-difference 338 /note= "encoded by CCG"
 FT /note= "encoded by CCG"
 FT US6111093-A.
 FT 29-AUG-2000.
 FT 28-OCT-1998; 98US-0181612.
 FT 01-DEC-1992; 92US-0983647.
 FT 25-FEB-1988; 88US-0160416.
 FT 13-JUL-1989; 89US-0379076.
 FT 23-MAR-1990; 90US-0498809.
 FT 13-JUL-1990; 90US-0535759.
 FT (GEHO) GEN HOSPITAL CORP.
 FT Stamenkovic I, Seed B;
 FT WPI: 2000-586382/55.
 FT N-PSDB; A50631.
 FT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 FT useful for immunodiagnosis and immunotherapy of immune-mediated
 FT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 FT diseases
 FT
 FT Example 10; Column 53-55; 75pb; English.
 FT
 FT The present sequence is that of a human macrophage specific FCRI,
 FT as deduced from CDNA clone p135 (see A50631), which was isolated
 FT from a cDNA library expressed in COS cells using a novel method of
 FT the invention designed to isolate CSA nucleic acids. The method is
 FT based upon transient expression of a CSA in eukaryotic cells and
 FT physical selection of cells expressing the antigen by adhesion to
 FT (panning on) an antibody-coated substrate such as a culture dish.
 FT CSA nucleic acids isolated by the method of the invention, and the
 FT proteins they encode, are useful for immunodiagnosis and
 FT immunotherapeutic applications, including the diagnosis and
 FT treatment of immune-mediated infections, diseases, and disorders in
 FT animals, including humans. These disorders include asthma,
 FT immune-complex disease, amyloidosis, parasitic diseases or multiple
 FT sclerosis. FCRI is a high affinity receptor for the Fc portion of
 FT IgG, normally located on the cell surfaces of macrophages. The
 FT ability to interfere with such bonding, or to cause it to occur on
 FT surfaces other than macrophages, is useful in therapy. A fusion
 FT protein of FCRI and a receptor ligand will be helpful to increase
 FT the potencies of antibodies in therapy.

```

Query Match      100.0%; Score 55; DB 21; Length 344;
Best Local Similarity 100.0%; P-Id. No. 0.0052;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 VNASVTSPLLE 12
        |||

```

```
Db      193 vlnasvtsplle 204

RESULT       6
ID          R20811
AC          R20811 standard; Protein; 374 AA.
XX
DT         21-MAY-1992 (first entry)
XX
DE        Human macrophage-specific FcRI receptor encoded by clone p135.
KW        Rapid immunoselection cloning technique; cell surface antigen;
KW        Immunodiagnosis; high affinity receptor.
OS        Homo sapiens.
PN         WO9201049-A.
PD         23-JAN-1992.
PF         15-JUL-1990;    90WO-US04986.
PR         13-JUL-1990; --   90US--0553759.
PP         (GEHO-) GEN HOSPITAL CORP.
PA         Seed B, Aruffo A, Amiot M;
PI         WPI: 1992-056664/07.
PX         N-PSSD; Q21178.
DR
XX
PT         New CD53 cell surface antigen and DNA encoding it - for
PT         immuno-therapy and diagnosis of hematopoietic neoplasms, etc.
PS         Example 10; Page 94d; 160pp; English.
CC
XX
CC        This amino acid sequence was predicted from the cDNA sequence of
CC        cDNA clone p135. It differs from the sequence predicted from
CC        clones p98/X2 and p90 (see Q21179 and Q21180, respectively) at
CC        position 25; p135 encodes a Ser residue and the other two clones
CC        predict a Thr residue. At position 58, p135 predicts Leu and p90
CC        predicts Val. Sequences predicted from all 3 clones show the
CC        typical features of a type I integral membrane protein and include
CC        a short hydrophobic signal sequence, a single 21-residue
CC        cytoplasmic domain, The extracellular portion contains six
CC        potential N-linked glycosylation sites and six Cys residues
CC        distributed among three C2 set Ig-related domains. A fusion protein
CC        of FcRI and a receptor ligand will be helpful to increase the
CC        potency of antibodies in therapy.
CX
SQ         Sequence     374 AA;
```

```
Query Match           100.0%; Score 55; DB 13; Length 374;
Best Local Similarity 100.0%; -Pred. No. 0.0058;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY              1 VLNASVTSPLLE 12
                  |||||
DB            193 VLNASVTSPILLE 204
```

```
RESULT       7
R22550
ID          R22550 standard; Protein; 374 AA.
XX
AC          R22550;
XX
DT         21-MAY-1992 (first entry)
```


DE Human macrophage-specific FcRI receptor encoded by clone p90.
 XX
 KM Rapid immunoselection cloning technique; cell surface antigen;
 KM Immunodiagnosis; high affinity receptor.
 XX
 OS Homo sapiens.
 XX
 PN W09201049-A.
 XX
 PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90MO-US04986.
 XX
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO-) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Amiot M;
 XX
 DR WPI: 1992-056864/07.
 DR N-PSDB; Q21180.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 PT Immuno-therapy and diagnosis of hematopoietic neoplasms, etc.
 XX
 PS Example 10; Page 94a; 160pp; English.
 XX
 CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p90. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC and clone p90 predicts a Thr residue. At position 58, p90 predicts
 CC Val while clones p135 and p98/X2 (see Q21179) both predict Leu.
 CC Sequences predicted from all 3 clones show the typical features of
 CC a type I integral membrane protein and include a short hydrophobic
 CC signal sequence, a single 21-residue hydrophobic membrane-spanning
 CC domain, and a short, highly charged cytoplasmic domain. The
 CC extracellular portion contains six potential N-linked glycosylation
 CC sites and six Cys residues distributed among three C2 set Ig-related
 CC domains. A fusion protein of FcRI and a receptor ligand will be
 CC helpful to increase the potency of antibodies in therapy.
 XX
 SQ Sequence 374 AA;

Query Match 100.0%; Score 55; DB 13; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNAVSTSPLE 12
 Db 193 VLNAVSTSPLE 204

RESULT 8
 R91438
 ID R91438 standard; Protein; 374 AA.
 XX
 AC R91438;
 XX
 DT 30-OCT-1996 (first entry)
 XX
 DE Human FcRI (cDNA clone p135 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FcRI; Fc receptor.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 25 /note="amino acid 25 is Thr in clone p90 and p98
 FT translated products"
 FT Misc-difference 58

FT /note="amino acid 58 is Val in p90 clone
 FT translated product."
 XX
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI: 1996-200279/20.
 DR N-PSDB; T14717.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX
 PS Example 10; Column 55-56; 79pp; English.
 XX
 CC The amino acid sequence (R91438) of human FcRI was detd. from a
 CC cDNA clone, p135 (T14717), obtd. from a cDNA library using an
 CC immunoselection cloning method. FcRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Another isolated cDNA clone, p90 (T14719), coded
 CC for a variant (W00859) of the FcRI sequence, and a third clone, p98
 CC (T14718), coded for an FcRI (R91439) having a different C-terminal
 CC sequence. A fusion protein of FcRI and a receptor ligand will
 CC be helpful in increasing the potency of antibodies in therapy.
 XX
 SQ Sequence 374 AA;

Query Match 100.0%; Score 55; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNAVSTSPLE 12
 Db 193 VLNAVSTSPLE 204

RESULT 9
 W00859
 ID W00859 standard; Protein; 374 AA.
 XX
 AC W00859;
 XX
 DT 30-OCT-1996 (first entry)
 XX
 DE Human FcRI (cDNA clone p90 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FcRI; Fc receptor.
 XX
 OS Homo sapiens.
 XX
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI; 1996-200279/20.
 DR N-PSDB; T14719.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX
 PS Example 10; Column 55-56; 79pp; English.
 XX
 CC The amino acid sequence (W00859) of human FCRI was detd. from a
 CC cDNA clone, p90 (T14719), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Other cDNA clones (see also T14717-18) coded for
 CC variants (F91438-39) of the FCRI sequence. A fusion protein of
 CC FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.
 CC
 SQ Sequence 374 AA;

Query Match 100.0%; Score 55; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNASTVSPLE 12
 Db 193 vlnastvsplle 204

RESULT 10
 ID W80448
 AC W80448; standard; Protein; 374 AA.

XX 07-JUN-1999 (first entry)
 DT
 XX Human Fc receptor I.
 DE
 XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 KM cloning.
 KM
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 2 /note= "encoded by TGG"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by CTG"
 FT Misc-difference 60 /note= "encoded by CCC"
 FT Misc-difference 77 /note= "encoded by AAT"
 FT Misc-difference 85 /note= "encoded by TCC"
 FT Misc-difference 99 /note= "encoded by CAA"
 FT Misc-difference 103 /note= "encoded by CCC"
 FT Misc-difference 141 /note= "encoded by GGC"
 FT Misc-difference 159

FT /note= "encoded by AAC"
 FT Misc-difference 171 /note= "encoded by ATG"
 FT Misc-difference 176 /note= "encoded by GTC"
 FT Misc-difference 256 /note= "encoded by GGC"
 FT
 XX US5830731-A.
 PN
 XX 03-NOV-1998.
 PD
 XX 21-MAY-1997; 97US-0861205.
 PF
 XX 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 21-MAY-1997; 97US-0861205.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 XX Aruffo A, Seed B;
 PI
 XX
 DR WPI; 1998-609251/51.
 DR N-PSDB; V63456.
 XX
 XX New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNAs, especially
 PT human lymphocyte antigenic sequences
 PT
 PS Example 10; Column 53-54; 75pp; English.

XX This is the amino acid sequence of human Fc receptor I (FCRI), as
 CC deduced from cDNA clone p135 (see V63456) isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from polyA RNA of cells
 CC from a single patient undergoing extracorporeal interferon-2
 CC induction therapy. Clones p90 (see X07372) and p98 (see X07373)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.

SQ Sequence 374 AA;

Query Match 100.0%; Score 55; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNASTVSPLE 12
 Db 193 vlnastvsplle 204

RESULT 11
 ID W97833
 AC W97833; standard; Protein; 374 AA.

XX 07-JUN-1999 (first entry)
 DT Human Fc receptor I.
 XX
 DE Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 KW cloning.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2 /note= "encoded by TGC"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by CTC"
 FT Misc-difference 60 /note= "encoded by CCC"
 FT Misc-difference 77 /note= "encoded by AAT"
 FT Misc-difference 85 /note= "encoded by TCC"
 FT Misc-difference 99 /note= "encoded by CAA"
 FT Misc-difference 103 /note= "encoded by CCC"
 FT Misc-difference 141 /note= "encoded by GGC"
 FT Misc-difference 159 /note= "encoded by AAC"
 FT Misc-difference 171 /note= "encoded by ATG"
 FT Misc-difference 176 /note= "encoded by GTC"
 FT Misc-difference 256 /note= "encoded by GGC"
 FT /note= "encoded by GGC"
 XX
 XX US5830731-A.
 PN
 XX
 PD 03-NOV-1998.
 XX
 PF 21-MAY-1997; 97US-0861205.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 21-MAY-1997; 97US-0861205.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI; 1998-609251/51.
 DR N-PSDB; X07372.
 XX
 XX New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNA(s), especially
 PT human lymphocyte antigenic sequences
 XX
 XX Example 10; Column 53-54; 75bp; English.
 CC This is the amino acid sequence of human Fc receptor I (FCRI), as
 CC deduced from cDNA clone p90 (see X07372) isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from polyA RNA of cells
 CC from a single patient undergoing extracorporeal interleukin-2
 CC induction therapy. Clones p135 (see V63456) and p98 (see X07373)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen

CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 XX
 XX Sequence 374 AA:
 SO
 CC
 CC Query Match 100.0%; Score 55; DB 19; Length 374;
 CC Best Local Similarity 100.0%; Pred. No. 0.0058;
 CC Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VLNAVSTSPLE 12
 DB 193 VLNAVSTSPLE 204
 W97834
 W97834 standard; Protein; 374 AA.
 XX
 AC W97834;
 XX
 DT 07-JUN-1999 (first entry)
 XX
 DE Human Fc receptor I.
 XX
 KW Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 KW cloning.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2 /note= "encoded by TGC"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by AAC"
 FT Misc-difference 60 /note= "encoded by ATG"
 FT Misc-difference 77 /note= "encoded by GTC"
 FT Misc-difference 85 /note= "encoded by CCC"
 FT Misc-difference 99 /note= "encoded by CAA"
 FT Misc-difference 103 /note= "encoded by CCC"
 FT Misc-difference 141 /note= "encoded by GGC"
 FT Misc-difference 159 /note= "encoded by AAC"
 FT Misc-difference 171 /note= "encoded by ATG"
 FT Misc-difference 176 /note= "encoded by GTC"
 FT Misc-difference 256 /note= "encoded by GGC"
 FT /note= "encoded by GGC"
 XX
 XX US5830731-A.
 PN

XX 03-NOV-1998.
 PD 21-MAY-1997; 97US-0861205.
 XX
 PF 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 21-MAY-1997; 97US-0861205.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI: 1998-609251/51.
 N-PSDB; X07373.
 XX
 XX New Cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNA(s), especially
 PT human lymphocyte antigenic sequences
 XX
 PS Example 10; Column 53-54; 75bp; English.
 XX
 XX This is the amino acid sequence of human Fc receptor I (FCRI), as
 CC deduced from cDNA clone p98/X2 (see X07373) isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from polyA RNA of cells
 CC from a single patient undergoing extracorporeal Interleukin-2
 CC induction therapy. Clones p135 (see V63456) and p90 (see X07372)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 CC
 SQ Sequence 374 AA;
 QY 1 VNASVTSPLE 12
 DB 193 vlnasvtsple 204
 Y6134
 Y6134 standard; Protein; 374 AA;
 Y6134;
 19-DEC-2000 (first entry)
 Human macrophage-specific FCRI.
 Macrophage; FCRI; cell surface antigen; human; immunoselection;
 panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 immune disorder; infection; asthma; immune-complex disease;
 amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 15 /note= "encoded by GCG"
 FT Misc-difference 38 /note= "encoded by ACC"
 FT Misc-difference 50 /note= "encoded by CCG"
 FT Misc-difference 51 /note= "encoded by ACC"
 FT Misc-difference 55 /note= "encoded by CAC"
 FT Misc-difference 56 /note= "encoded by TCC"
 FT Misc-difference 60 /note= "encoded by CCC"
 FT Misc-difference 64 /note= "encoded by CAG"
 FT Misc-difference 82 /note= "encoded by GAA"
 FT Misc-difference 116 /note= "encoded by GAA"
 FT Misc-difference 117 /note= "encoded by GCA"
 FT Misc-difference 122 /note= "encoded by TTC"
 FT Misc-difference 123 /note= "encoded by ACC"
 FT Misc-difference 126 /note= "encoded by CCG"
 FT Misc-difference 129 /note= "encoded by CAG"
 FT Misc-difference 134 /note= "encoded by AAT"
 FT Misc-difference 136 /note= "encoded by GTT"
 FT Misc-difference 139 /note= "encoded by CCA"
 FT Misc-difference 140 /note= "encoded by AAT"
 FT Misc-difference 213 /note= "encoded by CAA"
 FT Misc-difference 216 /note= "encoded by TTC"
 FT Misc-difference 220 /note= "encoded by CGT"
 FT Misc-difference 268 /note= "encoded by AAT"
 FT Misc-difference 305 /note= "encoded by CTC"
 FT Misc-difference 306 /note= "encoded by AAC"
 FT Misc-difference 332 /note= "encoded by GGT"
 US611093-A.
 29-AUG-2000.
 28-OCT-1998; 98US-0181612.
 01-DEC-1992; 92US-0983647.
 25-FEB-1988; 88US-0160416.
 13-JUL-1989; 89US-0379076.
 23-MAR-1990; 90US-0498809.
 13-JUL-1990; 90US-0553759.
 (GENO) GEN HOSPITAL CORP.
 Stamenkovic I, Seed B;

DR WPI: 2000-586382/55.
 DR N-PSDB; A50592.
 XX
 XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 PT useful for immunodiagnosis and immunotherapy of immune-mediated
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 PT diseases
 XX
 PS Example 10; Column 53-55; 75pp; English.
 XX
 XX The present sequence is that of a human macrophage specific FCRI,
 CC as deduced from cDNA clone p135 (see A50592), which was isolated
 CC from a cDNA library expressed in COS cells using a novel method of
 CC the invention designed to isolate CSA nucleic acids. The method is
 CC based upon transient expression of a CSA in eukaryotic cells and
 CC physical selection of cells expressing the antigen by adhesion to
 CC (panning on) an antibody-coated substrate such as a culture dish.
 CC CSA nucleic acids isolated by the method of the invention, and the
 CC proteins they encode, are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders in
 CC animals, including humans. These disorders include asthma,
 CC immune-complex disease, amyloidosis, parasitic diseases or multiple
 CC sclerosis. FCRI is a high affinity receptor for the Fc portion of
 CC IgG, normally located on the cell surfaces of macrophages. The
 CC ability to interfere with such bonding, or to cause it to occur on
 CC surfaces other than macrophages, is useful in therapy. A fusion
 CC protein of FCRI and a receptor ligand will be helpful to increase
 CC the potencies of antibodies in therapy.
 CC
 CC Sequence 374 AA;
 SO
 Query Match 100.0%; Score 55; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VNASTVSPLE 12
 Db 193 vlnastvspile 204
 RESULT 14
 Y96226
 ID Y96226 standard; Protein: 374 AA.
 AC Y96226;
 XX
 DT 11-SEP-2000 (first entry)
 XX
 DE Human high affinity Fc receptor; Fc gamma mAb.
 XX
 XX Human; high affinity Fc receptor; Fc gamma mAb; immunoglobulin;
 KM infection; immune response; CD64; monocytic; macrophage; neutrophil;
 KM eosinophil; HIV; IgG; immunosuppressive; antirheumatic; cytostatic;
 KM antiinflammatory; AIDS; SLE; multiple myeloma; rheumatoid arthritis;
 KM systemic lupus erythematosus; tumour.
 XX
 OS Homo sapiens.
 XX
 PN EP1006183-A1.
 XX
 PD 07-JUN-2000.
 XX
 PF 03-DEC-1998; 98EP-0122969.
 XX
 PR 03-DEC-1998; 98EP-0122969.
 XX
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX
 DR WPI: 2000-367968/32.
 DR N-PSDB; A27466.
 XX

PT Novel Fc receptor lacking transmembrane domains, a signal peptide, and
 PT glycosylation, useful for diagnosing and treating immune disorders and
 PT cancer
 XX
 XX Disclosure; Page 26-28; 60pp; English.
 PS
 XX
 XX The present sequence is the human high affinity Fc receptor, Fc gamma mAb.
 CC Fc gamma mAb is also known as CD64. Fc receptors play an important
 CC role in defending the body against infections. First, pathogens are
 CC opsonised by serum immunoglobulins. The resulting complex then binds to
 CC cells expressing Fc receptors. Fc gamma mAb molecules are expressed by
 CC monocytes and macrophages, but expression can also be induced on
 CC neutrophils and eosinophils. Upon Fc receptor activation, immune effector
 CC pathways are activated, leading to immune response. The present sequence
 CC may be modified to produce recombinant versions. The recombinant Fc
 CC receptor consist only of the extracellular portion of the receptor and
 CC are not glycosylated i.e. they do not have transmembrane domains or
 CC signal peptides. The recombinant proteins may be used in immunoassays to
 CC determine the immune status of patients with chronic diseases of the
 CC immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple
 CC myeloma (MM), or rheumatoid arthritis. In addition, pharmaceutical
 CC compositions containing recombinant proteins may be used to treat or
 CC prevent autoimmune diseases, allergies or tumours, especially AIDS,
 CC rheumatoid arthritis or MM.
 CC
 CC Sequence 374 AA;
 SO
 Query Match 100.0%; Score 55; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VNASTVSPLE 12
 Db 193 vlnastvspile 204
 RESULT 15
 B43683
 ID B43683 standard; Protein: 399 AA.
 AC B43683;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1128.
 XX
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KM diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
 KM antidiabetic; antisthmatic; antirheumatic; antiarthritic; antiviral;
 KM dermatological; antihypertensive; antiallergic; antibacterial; cardiant;
 KM vasotropic; antiproliferative; thrombolytic; coagulant; neurotropic;
 KM immune disorder; hematopoietic cell disorder; autoimmune disorder;
 KM allergic reaction; graft versus host disease; organ rejection;
 KM haemostatic; thrombolytic; cardiovascular disorder; infection;
 KM neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2000-587533/55.
 XX

DR N-PSDB: C77892.

XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer -PS
XX Claim 11; Page 1739-1740; 2352pp; English.XX
CC C77607 to C78448 encode the human cancer associated proteins given in
CC B43398 to B44239. The proteins can have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic;
CC antiasthmatic; antirheumatic; antiarthritis; antiinflammatory;
CC antithyroid; antiallergic; antibacterial; antiviral; dermatological;
CC neuroprotective; cardiac; thrombolytic; coagulant; nootropic;
CC vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and
CC polypeptides can be used for preventing, treating or ameliorating medical
CC conditions and diagnosing pathological conditions. Polynucleotides,
CC polypeptides, antibodies, agonists and antagonists from the present
CC invention may be used to treat immune disorders by activating or
CC inhibiting the proliferation, differentiation or mobilisation of immune
CC cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC allergic reactions, graft versus host disease and organ rejection,
CC modulate haemostatic or thrombolytic activity, modulate inflammation,
CC cancers, cardiovascular disorders, neurological disease and bacterial or
CC viral infections. The peptides, nucleotides, antibodies, agonists and
CC antagonists may be also be used in drug screens. C78449 to C78457 and
CC B44240 represent sequences used in the exemplification of the present
XX invention.

SQ Sequence 399 AA;

Query Match 100.0%; Score 55; DB 21; Length 399;

Best Local Similarity 100.0%; Pred. No. 0.0063;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VINASTSPLE 12

|||||

Db 218 vinastsp1e 229

Search completed: June 4, 2001, 12:13:26
Job time: 199 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:10:47 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-21

Perfect score: 55
Sequence: 1 VLNSVTSPLLE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6ackfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	61.8	91	2	US-08-332-562A-94
2	34	61.8	3218	1	US-08-764-100-27
3	33	60.0	131	2	US-08-721-488-8
4	33	60.0	350	6	5352575-7
5	33	60.0	788	1	US-08-572-225-1
6	32	58.2	661	1	US-08-232-538-12
7	32	58.2	687	2	US-08-786-164-6
8	32	58.2	687	1	US-08-232-538-6
9	32	58.2	687	2	US-08-786-164-6
10	32	58.2	758	2	US-08-874-678-1
11	32	58.2	758	3	US-08-643-839-1
12	32	58.2	780	1	US-08-332-538-14
13	32	58.2	780	2	US-08-786-164-14
14	32	58.2	1311	1	US-08-340-011-5
15	32	58.2	1311	3	US-08-901-710-5
16	32	58.2	1338	3	US-08-750-141A-3
17	32	58.2	1362	2	US-08-874-678-33
18	32	58.2	1362	2	US-08-643-839-33
19	31	56.4	20	2	US-08-564-972-20
20	31	56.4	193	2	US-08-564-972-8
21	31	56.4	194	1	US-08-148-058A-27
22	31	56.4	194	1	US-08-148-058A-29
23	31	56.4	194	1	US-08-478-042-27
24	31	56.4	194	1	US-08-478-042-29
25	31	56.4	194	1	US-08-645-215-27
26	31	56.4	194	2	US-08-645-215-29
27	31	56.4	194	2	US-08-466-604-27

28	31	56.4	194	2	US-08-466-604-29	Sequence 29, Appl
29	31	56.4	247	1	US-08-324-977-44	Sequence 44, Appl
30	31	56.4	247	2	US-08-384-616-44	Sequence 44, Appl
31	31	56.4	247	2	US-08-304-686A-44	Sequence 44, Appl
32	31	56.4	285	4	US-09-141-821-4	Sequence 4, Appl1
33	31	56.4	382	1	US-08-415-818-7	Sequence 7, Appl1
34	31	56.4	382	2	US-08-894-236-7	Sequence 7, Appl1
35	31	56.4	382	2	US-08-555-268A-14	Sequence 14, Appl1
36	31	56.4	382	5	PCT-US86-01444-7	Sequence 7, Appl1
37	31	56.4	390	4	US-09-036-987A-8	Sequence 8, Appl1
38	31	56.4	535	2	US-08-564-972-1	Sequence 1, Appl1
39	31	56.4	572	5	PCT-US91-08177-11	Sequence 11, Appl1
40	31	56.4	572	5	PCT-US91-08177-19	Sequence 19, Appl1
41	31	56.4	617	1	US-08-279-700-2	Sequence 2, Appl1
42	31	56.4	617	1	US-08-279-700-4	Sequence 4, Appl1
43	31	56.4	617	1	US-08-279-700-6	Sequence 6, Appl1
44	31	56.4	617	1	US-08-279-700-8	Sequence 8, Appl1
45	31	56.4	617	1	US-08-279-700-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1
US-08-332-562A-94
Sequence 94, Application US/08332562A
Patent No. 5985599
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HOLETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-94
Query Match 61.8%, Score 34, DB 2, Length 91;

Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VNASTVSPLE 12
11111111
Db 5 VNASTVSPFPE 16

RESULT 2

US-08-764-100-27
Sequence 27, Application US/08764100
Patent No. 5773700

GENERAL INFORMATION:

APPLICANT: van Grinsven J., Martinus Q.
APPLICANT: De Haan, Petrus T.
APPLICANT: Gieken L., Johannes J.
APPLICANT: Peters, Dirk
APPLICANT: Goldbach, Robert W.
TITLE OF INVENTION: Improvements in or Relating to organic
TITLE OF INVENTION: Compounds
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,100
FILING DATE: 06-DEC-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,064
FILING DATE:

APPLICATION NUMBER: US 08/032,235
FILING DATE: 17-MAR-1993
APPLICATION NUMBER: GB 9206016.9
FILING DATE: 19-MAR-1992

ATTORNEY/AGENT INFORMATION:
NAME: No. 5773700-1's, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER: 137-1061

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 354-3592
TELEFAX: (415) 857-1125
INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 3218 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: protein
US-08-764-100-27

Query Match 61.8%; Score 34; DB 1; Length 3218;
Best Local Similarity 63.6%; Pred. No. 8.7e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNASTVSPLE 12
11111111
Db 2230 LNASTVSPLE 2240

RESULT 3

US-08-721-488-8
Sequence 8, Application US/08721488

Patent No. 5965388

GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: Lavallie, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Bowman, Michael
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,488
FILING DATE:

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.

REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-5851
TELEFAX: (617) 498-8224

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 131 amino acids
TYPE: amino acid
STRANDEDNESS:

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-721-488-8

Query Match 60.0%; Score 33; DB 2; Length 131;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VNASTVSPLE 12
11111111
Db 13 IFDAVXTPTLE 24

RESULT 4

5352575-7
Patent No. 5352575
APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.
TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/513,282
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 100,817
FILING DATE: 29-JUN-1987
APPLICATION NUMBER: 886,260
FILING DATE: 16-JUL-1986
APPLICATION NUMBER: 784,787
FILING DATE: 04-OCT-1985
APPLICATION NUMBER: 801,799
FILING DATE: 26-NOV-1985
APPLICATION NUMBER: 844,113

FILING DATE: 26-MAR-1986
SEQ ID NO: 7
LENGTH: 350
5352575-7

Query Match 60.0%; Score 33; DB 6; Length 350;
Best Local Similarity 72.7%; Pred. No. 97;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASVSP.L 11
| | | | | : |
Db 260 VLNASVSP.L 270

RESULT 5
US-08-572-225-1
Sequence 1, Application US/08572225
Patent No. 5807981
GENERAL INFORMATION:
APPLICANT: PROCTOP, Darwin J.
APPLICANT: HOJIMA, Yoshio
APPLICANT: LI, Shi-Wu
APPLICANT: SIERON, Alexander
APPLICANT: BRENNER, Mitch
TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND ITS USE FOR
TITLE OF INVENTION: DRUG DEVELOPMENT FOR THE TREATMENT OF DISEASE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/572,225
FILING DATE: 13-DEC-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 8389-031
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-572-225-1

Query Match 60.0%; Score 33; DB 1; Length 788;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LNASVSP 9
| | | | | : |
Db 400 LNASVSP 407

RESULT 6
US-08-232-538-12

Sequence 12, Application US/08232538
Patent No. 5712380
GENERAL INFORMATION:
APPLICANT: Thomas, Kenneth A.
APPLICANT: KENDALL, Richard L.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000 126 E Lincoln Avenue
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,538
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wallen, John W. III
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 188881A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-232-538-12

Query Match 58.2%; Score 32; DB 1; Length 661;
Best Local Similarity 60.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASVSP.L 10
| | | | | : |
Db 223 VLNASVSP.L 232

RESULT 7
US-08-786-164-12
Sequence 12, Application US/08786164
Patent No. 5861484
GENERAL INFORMATION:
APPLICANT: THOMAS, KENNETH A.
APPLICANT: KENDALL, RICHARD L.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL
TITLE OF INVENTION: CELL GROWTH FACTOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/786,164
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mark Hand, J
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18888DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-786-164-12

Query Match
Best Local Similarity 58.2%; Score 32; DB 2; Length 661;
Best Local Similarity 60.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASTV SPL 10
||| : |||
Db 223 VLNCTATTPL 232

RESULT 8
US-08-232-538-6
Sequence 6, Application US/08232538,
Patent No. 5712380
GENERAL INFORMATION:
APPLICANT: Thomas, Kenneth A.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000 126 E Lincoln Avenue
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,538
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wallen, John W.III
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 188881A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4720
TELEFAX: (908) 594-3905
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-232-538-6

Query Match
Best Local Similarity 58.2%; Score 32; DB 1; Length 687;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASTV SPL 10
||| : |||
Db 249 VLNCTATTPL 258

RESULT 9
US-08-786-164-6
Sequence 6, Application US/08786164
Patent No. 5861484
GENERAL INFORMATION:
APPLICANT: THOMAS, KENNETH A.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL
CELL GROWTH FACTOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,164
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mark Hand, J
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18888DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-786-164-6

Query Match
Best Local Similarity 58.2%; Score 32; DB 2; Length 687;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASTV SPL 10
||| : |||
Db 249 VLNCTATTPL 258

RESULT 10
US-08-874-678-1
Sequence 1, Application US/08874678
Patent No. 5952199
GENERAL INFORMATION:

APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Presta, Leonard
APPLICANT: Ferrara, Napoleone
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR
TITLE OF INVENTION: ACTIVITY, THEIR USES AND PROCESSES FOR THEIR PRODUCTION
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,678
CLASSIFICATION: 435
FILING DATE: HEREWITH
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/643,839
FILING DATE: 07-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63291-1/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 758 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-874-678-1

Query Match 58.2%; Score 32; DB 2; Length 758;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLNASVTSPL 10
DB 249 VLNCTATPL 258

RESULT 11
US-08-643-839-1
Sequence 1, Application US/08643839
Patent No. 6100071
GENERAL INFORMATION:
APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Presta, Leonard
APPLICANT: Ferrara, Napoleone
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/643,839
FILING DATE: 07-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63291/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 758 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-643-839-1

Query Match 58.2%; Score 32; DB 3; Length 758;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLNASVTSPL 10
DB 249 VLNCTATPL 258

RESULT 12
US-08-232-538-14
Sequence 14, Application US/08232538
Patent No. 5712380
GENERAL INFORMATION:
APPLICANT: Thomas, Kenneth A.
APPLICANT: Kendall, Richard L.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000 126 E Lincoln Avenue
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,538
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wallen, John W. III
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 188881A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 780 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-232-538-14

Query Match
Best Local Similarity 58.2%; Score 32; DB 1; Length 780;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLNASTVSTPL 10
DB 249 VLNCTATPTL 258

RESULT 13
US-08-786-164-14
Sequence 14, Application US/08786164
Patent No. 5861484

GENERAL INFORMATION:
APPLICANT: THOMAS, KENNETH A.
TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL
NUMBER OF INVENTIONS: CELL GROWTH FACTOR
CORRESPONDENCE ADDRESS:
ADDRESS: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786/164
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
NAME: Mark Hand, J
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18888DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX:

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 780 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-786-164-14

Query Match
Best Local Similarity 58.2%; Score 32; DB 2; Length 780;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLNASTVSTPL 10
DB 249 VLNCTATPTL 258

RESULT 14
US-08-340-011-5

Sequence 5, Application US/08340011
Patent No. 5776755

GENERAL INFORMATION:
APPLICANT: Aitalo, et al.
TITLE OF INVENTION: FLT4, A NOVEL RECEPTOR TYROSIN KINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340/011
FILING DATE:

CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/959,951
FILING DATE: 09-OCT-1992
ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Thomas C.
REGISTRATION NUMBER: 36,989
REFERENCE/DOCKET NUMBER: 32267
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-340-011-5

Query Match
Best Local Similarity 58.2%; Score 32; DB 1; Length 1311;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLNASTVSTPL 10
DB 256 VLNCTATPTL 265

RESULT 15
US-08-901-710-5
Sequence 5, Application US/08901710
Patent No. 6107046

GENERAL INFORMATION:
APPLICANT: Aitalo, Karl
APPLICANT: Aprelikova, Olga
APPLICANT: Pajusola, Katari
APPLICANT: Armstrong, Elin
APPLICANT: Korhonen, Jaana
APPLICANT: Kaipainen, Arja
APPLICANT: Mäkeläinen, Marja-Terttu
TITLE OF INVENTION: FLT4, A RECEPTOR TYROSINE KINASE, AND USES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America

ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/901,710
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,754
FILING DATE: 09-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/959,951
FILING DATE: 09-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/33824
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-901-710-5

Query Match 58.2%; Score 32; DB 3; Length 1311;
Best Local Similarity 60.0%; Pred. NO. 7.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 VLNASVTSPL 10
||| : |||
Db 256 VLNCTATFPL 265

Search completed: June 4, 2001, 12:14:28
Job time: 221 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:11:07 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-21

Perfect score: 55
Sequence: 1 VLNSVTSPLLE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR-67:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	344	2 A41357	Fc gamma (19g) rec
2	55	100.0	374	1 A39878	Fc gamma (19g) rec
3	38	69.1	347	1 S51120	alcohol dehydrogen
4	37	67.3	175	2 A70824	hypothetical glyci
5	37	67.3	361	1 D69005	hypothetical prote
6	37	67.3	1118	2 T28426	probable DNA-direc
7	35	63.6	240	1 S04645	transcription regu
8	35	63.6	347	1 A44245	alcohol dehydrogen
9	35	63.6	422	2 H64489	hypothetical prote
10	35	63.6	591	2 T48596	ankyrin-like prote
11	35	63.6	632	2 C71327	probable phosphog
12	35	63.6	761	1 S61642	translation elonga
13	35	63.6	1223	2 S62811	PHOS5 protein - ye
14	34	61.8	256	1 C69304	probable enoyl-CoA
15	34	61.8	336	2 I48471	Fc gamma (19g) rec
16	34	61.8	367	2 E70347	cell division prot
17	34	61.8	393	2 T51092	cell division prot
18	34	61.8	402	2 T30667	hypothetical prote
19	34	61.8	404	2 A46480	Fc gamma (19g) rec
20	34	61.8	423	2 T33667	hypothetical prote
21	34	61.8	538	2 S67766	RNA-export mediat
22	34	61.8	570	2 B70204	phosphomannomase
23	34	61.8	1107	2 S67381	tubulin-folding co
24	33	60.0	76	2 T25127	hypothetical prote
25	33	60.0	124	2 S03231	hypothetical prote
26	33	60.0	145	2 T13542	hypothetical prote
27	33	60.0	156	2 C72506	hypothetical prote
28	33	60.0	201	2 T49213	hypothetical prote
29	33	60.0	209	2 A69864	conserved hypothet

30	33	60.0	251	2 A75482	oxidoreductase, sh
31	33	60.0	288	2 T41112	hypothetical prote
32	33	60.0	350	1 VGB563	glycoprotein gp63
33	33	60.0	379	1 VGC4289	cell division prot
34	33	60.0	405	2 G70829	probable molybdenu
35	33	60.0	418	2 H82209	conserved hypothet
36	33	60.0	428	4 S46932	hypothetical cell-
37	33	60.0	430	2 S77393	cell division prot
38	33	60.0	524	2 T06021	hypothetical prote
39	33	60.0	524	2 A82580	polyvinylalcohol d
40	33	60.0	615	2 S77332	NADH dehydrogenase
41	33	60.0	730	1 BMH01	procollagen C-endo
42	33	60.0	823	1 A58788	procollagen C-endo
43	33	60.0	986	1 B58788	procollagen C-endo
44	33	60.0	991	2 I49540	procollagen C-endo
45	33	60.0	1082	2 T43990	hypothetical prote

ALIGNMENTS

RESULT 1
Fc gamma (19g) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C>Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000
C/Accession: A41357; S03019
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A>Title: Isolation and expression of functional high-affinity Fc receptor complementa
A:Reference number: A41357; MID:89100284
A:Accession: A41357
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: GB:X14355; GB:M21090; NID:g31333; PIDN:CAA32536.1; PID:g31334
R:Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A>Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (C/Accession: S03019; MID:89098339)
A:Reference number: S03018; MID:89098339
A:Accession: S03019
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: EMBL:X14355; NID:g31333; PIDN:CAA32536.1; PID:g31334
A>Note: the authors translated the codon ACT for residue 25 as Ala
C:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
F:117-170/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 55; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNSVTSPLLE 12
DB 193 VLNSVTSPLLE 204

RESULT 2
Fc gamma (19g) receptor I-A (high affinity) precursor - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C>Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C/Accession: A39878; I70304; B41357; S03018; I57525
R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
J. Biol. Chem. 266, 13449-13455, 1991
A>Title: Gene organization of the human high affinity receptor for IgG, Fc gamma1 (CD
A:Reference number: A39878; MID:91302383
A:Accession: A39878
A:Molecule type: DNA
A:Residues: 1-374 <VAN>

A:Cross-references: GB:M63830; GB:M63835; NID:g180277; PIDN:AAA35678.1; PID:g180279
 R.Porges, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: 155577; MUID:93055454
 A:Accession: 170304
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:L03418; NID:g184840; PIDN:AAA36049.1; PID:g292169
 R.Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: AA1357; MUID:89100284
 A:Accession: BA1357
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'T', '339-374 <ALL>
 A:Cross-references: GB:X14356; GB:M21091; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R.Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR)
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'T', '339-374 <AL2>
 A:Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R.Perez, C.; Mletzger, J.; Benesh, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A:Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
 1sm.
 A:Reference number: 157525; MUID:93204964
 A:Accession: 157525
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RE2>
 A:Cross-references: GB:S57204; NID:g298692; PIDN:AAJ3887.1; PID:g4261587
 C:Genetics:
 A:Gene: GDB:FCGR1A; CD64
 A:Cross-references: GB:135911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane F
 F:1-15/Domain: signal sequence status predicted <SIG>
 F:16-292/Domain: extracellular status predicted <EXT>
 F:117-170/Domain: immunoglobulin homology <IM2>
 F:293-313/Domain: transmembrane status predicted <TM2>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 55; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0042;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNASVTSPLLE 12
 DB 193 VNASVTSPLLE 204

RESULT 3
 S51120
 alcohol dehydrogenase (EC 1.1.1.1) - Sulfolobus sp.
 C:Species: Sulfolobus sp.
 C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 11-Jun-1999
 C:Accession: S51120
 R.Cannio, R.; Fiorentino, G.; Carpinelli, P.; Rossi, M.; Bartolucci, S.
 submitted to the EMBL Data Library, January 1995
 A:Description: Cloning and overexpression in Escherichia coli of the genes encoding NAD-
 A:Reference number: S51120
 A:Accession: S51120
 A:Molecule type: DNA
 A:Residues: 1-347 <CAN>
 A:Cross-references: EMBL:Z47543; NID:g623347; PIDN:CAA87591.1; PID:g623348
 A:Experimental source: strain RC3

C:Complex: homodimer
 C:Function:
 A:Description: catalyzes the oxidation of primary and secondary alcohols to aldehydes
 A:Pathway: alcohol degradation
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
 C:Keywords: alcohol metabolism; heat-stable protein; homodimer; metalloprotein; methyl
 F:23-337/Domain: long-chain alcohol dehydrogenase homology <LADH>
 F:174-203/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:11,213/Modified site: N6-methyllysine (Lys) #status predicted
 F:38,68,154/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted
 F:98,101,104,112/Binding site: zinc, noncatalytic (Glu, Cys, Cys, Cys) #status predic

Query Match 69.1%; Score 38; DB 1; Length 347;
 Best Local Similarity 66.7%; Pred. No. 9.2;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VNASVTSPLLE 12
 DB 220 VNASVDTPLAE 231

RESULT 4
 A70824
 hypothetical glycine-rich protein RV0742 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: A70824
 R.Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 R.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A:Reference number: A70500; MUID:98295987
 A:Accession: A70824
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-175 <CO>
 A:Cross-references: GB:AL021958; GB:AL124346; NID:g3261536; PIDN:CAA17509.1; PID:el25
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV0742

Query Match 67.3%; Score 37; DB 2; Length 175;
 Best Local Similarity 77.8%; Pred. No. 6.5;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 ASVTSPLLE 12
 DB 96 ASVTSPLLD 104

RESULT 5
 D69005
 hypothetical protein MTH1037 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: D69005
 R.Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T
 R.; Qiu, D.; Spadafora, R.; Vitale, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
 K.L.S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514
 A:Accession: D69005
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-361 <MTH>
 A:Cross-references: GB:AE000876; GB:AE000666; NID:g2622140; PIDN:AA85533.1; PID:g262
 A:Experimental source: strain Delta H
 C:Genetics:

A:Gene: MTH1037
C:Superfamily: hypothetical protein MTH778

Query Match 67.3%; Score 37; DB 1; Length 361;
Best Local Similarity 72.7%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASVTSPL 11
DB 172 VLKNSVLSPL 182

RESULT 6
T28426
probable DNA-directed DNA polymerase (EC 2.7.7.7) - Pichia kluyveri mitochondrion linear

C:Species: mitochondrion Pichia kluyveri
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T28426

R:Nosek, J.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z20485
A:Accession: T28426
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1118 <NOS>
A:Cross-references: EMBL:Y11606; NID:e1033579; PID:e321611; PIDN:CAA72340.1
C:Genetics:
A:Genome: mitochondrion linear plasmid
C:Keywords: mitochondrion; nucleotidyltransferase

Query Match 67.3%; Score 37; DB 2; Length 1118;
Best Local Similarity 71.4%; Pred. No. 57;
Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 1 VLNASV--TSPLE 12
DB 551 VLNASLKRTPLE 564

RESULT 7
S04645
transcription regulator farr, fatty acyl-responsive - Escherichia coli

N:Alternate names: gp30 protein
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S04645; S50896; A64809; PC4138
R:Buck, D.; Guest, J.R.
Biochem. J. 260, 737-747, 1989

A:Title: Overexpression and site-directed mutagenesis of the succinyl-CoA synthetase of
A:Reference number: S04645; MUID:89350876
A:Accession: S04645
A:Molecule type: DNA
A:Residues: 1-240 <BUC>
A:Cross-references: EMBL:X15790; NID:g41518; PIDN:CAA33791.1; PID:g41519
R:Quail, M.A.; Dempsey, C.E.; Guest, J.R.
FEBS Lett. 336, 183-187, 1994

A:Title: Identification of a fatty acyl responsive regulator (Farr) in Escherichia coli.
A:Reference number: S50896; MUID:95104410
A:Accession: S50896
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <QDA>
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: A64809
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-240 <BLAT>

A:Cross-references: GB:AE000176; GB:U00096; NID:g1786947; PIDN:AACT3824.1; PID:g17869
A:Experimental source: strain K-12, substrain MG1655
R:Utsunli, R.; Horie, T.; Katoh, A.; Kaino, Y.; Tanabe, H.; Noda, M.
Biosci. Biotechnol. Biochem. 60, 309-315, 1996

A:Title: Isolation and characterization of the heat-responsive genes in Escherichia c
A:Reference number: JCA598; MUID:97076625
A:Accession: PC4138
A:Molecule type: DNA
A:Residues: 1-40 <UTS>

A:Cross-references: DDBJ:D64014
A:Experimental source: R01012
C:Genetics:
A:Gene: farr; gp30
A:Map position: 16.8 min

C:Function: probably a transcription regulator of the succinyl-CoA synthase operon
A:Description: transcription regulator gntR
C:Superfamily: transcription regulator gntR
C:Keywords: DNA binding; transcription regulation

Query Match 63.6%; Score 35; DB 1; Length 240;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VLNASVTSPL 12
DB 192 LKNISOTRPLE 203

RESULT 8
A44245
alcohol dehydrogenase (EC 1.1.1.1) - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Jun-1999
C:Accession: A44245
R:Amendola, S.; Raita, C.A.; Caruso, C.; Camardella, L.; D'Auria, S.; De Rosa, M.; Ro
Biochemistry 31, 12514-12523, 1992

A:Title: Thermotable NAD(+) dependent alcohol dehydrogenase from Sulfolobus solfatar
A:Reference number: A44245; MUID:93099126
A:Accession: A44245
A:Molecule type: DNA
A:Residues: 1-347 <ANM>

A:Cross-references: GB:S51211; NID:g261968; PIDN:AB24546.1; PID:g261969
A:Experimental source: strain DSM1617
A:Note: sequence extracted from NCBI backbone (NCBI:120928, NCBI:120938)
A:Note: part of this sequence was confirmed by protein sequencing
C:Function: homodimer

A:Description: catalyzes the oxidation of primary and secondary alcohols to aldehydes
A:Pathway: alcohol degradation
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
C:Keywords: alcohol metabolism; heat-stable protein; homodimer; metalloprotein; methy
F:23-337/Domain: long-chain alcohol dehydrogenase homology <LADH>
F:174-203/Region: beta-alpha-beta NAD nucleotide-binding fold
F:11,213/Modified site: N6-methyllysine (Lys) (partial) #status experimental
F:38,68,154/binding site: zinc, catalytic (Cys, His, Cys) #status predicted
F:98,101,104,112/binding site: zinc, noncatalytic (Glu, Cys, Cys, Cys) #status predic

Query Match 63.6%; Score 35; DB 1; Length 347;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VLNASVTSPL 12
DB 220 VLNASMODPLAE 231

RESULT 9
H64489
hypothetical protein MJ1521 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Sep-1998

C:Accession: H64489
 R:Built, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reisch, C.I.; Overbeek, R.; Kikness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodok, A.; Iron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, J.M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999
 A:Accession: H64489
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-422 <BUD>
 A:Cross-References: GB:U67593; GB:L77117; NID:q1592149; PID:q1500411; TIGR:MJ1521
 C:Genetics:
 A:Map position: FOR1498907-1500175
 A:Start codon: GTG
 C:Superfamily: hypothetical protein MJ0057

Query Match 63.6%; Score 35; DB 2; Length 422;
 Best Local Similarity 58.3%; Pred. No. 45;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 1 VNASTSPL 12
 DB 292 LNCASISPL 303

RESULT 10
 T48596
 ankryrin-like protein - Arabidopsis thaliana
 N:Alternate names: protein F18022.20
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48596
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Meyers, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24493
 A:Accession: T48596
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-591 <BEV>
 A:Cross-References: EMBL:DB
 A:Experimental source: cultivar Columbia; BAC clone F18022
 C:Genetics:
 A:Map position: 5
 A:Introns: 32/3; 78/3
 A:Note: F18022.20

Query Match 63.6%; Score 35; DB 2; Length 591;
 Best Local Similarity 88.9%; Pred. No. 67;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 3 NASVTSPL 11
 DB 281 NASVTSPL 289

RESULT 11
 C71327
 probable phosphoglucosyltransferase - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 C:Accession: C71327
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Weidman, J.; Richardson, D.; Howell, J.K.; Childamaram, M.; Utterback, T.; MCD
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770
 A:Accession: C71327
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-632 <COL>
 A:Cross-References: GB:AE001219; GB:AE000520; NID:q3322693; PIDN:AA65401.1; PID:q332
 C:Genetics:
 A:Experimental source: strain Nichols
 A:Gene: TP0413

Query Match 63.6%; Score 35; DB 2; Length 632;
 Best Local Similarity 70.0%; Pred. No. 73;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 2 LNASTSPL 11
 DB 182 LNASTSPL 191

RESULT 12
 S61642
 translation elongation factor EF-G, mitochondrial - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein L2195; protein YLR069c
 C:Species: Saccharomyces cerevisiae
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: S61642; S20179; S64897; S64901; S17025
 R:Ureastarazu, L.A.
 submitted to the EMBL Data Library, December 1995
 A:Reference number: S61618
 A:Accession: S61642
 A:Molecule type: DNA
 A:Residues: 1-761 <URR>
 A:Cross-References: EMBL:X94607; NID:q1181264; PIDN:CAA64315.1; PID:q1181289
 R:Yambur, A.; Ackerman, S.H.; Tzagoloff, A.
 Eur. J. Biochem. 201, 643-652, 1991
 A:Title: Mitochondrial translational-initiation and elongation factors in Saccharomyc
 A:Reference number: S20177; MUID:92037620
 A:Accession: S20179
 A:Molecule type: DNA
 A:Residues: 1-65; 'I', 67-232, 'V', 234-477, 'S', 479-628, 'P', 630-761 <VAM>
 A:Cross-References: EMBL:X58378; NID:q3917; PIDN:CAA41267.1; PID:q3918
 R:Andre, B.; Ureastarazu, L.A.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64872
 A:Accession: S64897
 A:Molecule type: DNA
 A:Residues: 1-761 <AND>
 A:Cross-References: EMBL:Z73241; NID:q1360421; PIDN:CAA97626.1; PID:q1360422; GSPDB:G
 A:Experimental source: strain S288C
 R:Pohl, T.M.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64899
 A:Accession: S64901
 A:Molecule type: DNA
 A:Residues: 1-761 <POH>
 A:Cross-References: EMBL:Z73241; NID:q1360421; PIDN:CAA97626.1; PID:q1360422; GSPDB:G
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:MEF1; MIPS:YLR069c
 A:Cross-References: SGD:S0004059; MIPS:YLR069c
 A:Map position: 12R
 C:Superfamily: translation elongation factor G; translation elongation factor Tu homo
 C:Keywords: GTP binding; mitochondrial; P-loop; protein biosynthesis
 F:71-205/Domain: translation elongation factor Tu homology <ETU>
 F:77-84/Region: nucleotide-binding motif A (P-loop)
 F:202-205/Region: GTP-binding NKXD motif
 F:327-329/Region: GTP-binding SAK/L motif

Query Match 63.6%; Score 35; DB 1; Length 761;
 Best Local Similarity 66.7%; Pred. No. 91;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VNASTSPL 9

Db 668 INNAVSTSP 676

RESULT 13

S62011

PH085 protein - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein LPH16C; protein YPL110C

C:Species: *Saccharomyces cerevisiae*

C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 05-Nov-1999

C:Accession: S62011

R:Schlenstedt, G.; Silver, P. A.

A:Submitted to the EMBL Data Library, December 1995

A:Reference number: S61996

A:Accession: S62011

A:Molecule type: DNA

A:Residues: 1-1223 <SCH>

A:Cross-references: EMBL:U43503; NID:g1163087; PIDN:AAB68251.1; PID:g1163103; GSPDB:GNOC

C:Genetics:

A:Gene: PH085; MIPS:YPL110C

A:Map position: 16L

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

C:Keywords: transmembrane protein

F:577-593/Domain: transmembrane #status predicted <TM1>

F:1110-1126/Domain: transmembrane #status predicted <TM2>

Query Match

Best Local Similarity 63.6%; Score 35; DB 2; Length 1223;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNAVSTSP.LL 11

Db 821 ILNGSVTPIL.831

RESULT 14

C69304

probable enoyl-CoA hydratase (EC 4.2.1.17) fad-1 - *Archaeoglobus fulgidus*C:Species: *Archaeoglobus fulgidus*

C:Date: 28-May-1999 #sequence_revision 28-May-1999 #text_change 18-Jun-1999

C:Accession: C69304

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MID:98049343

A:Accession: C69304

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-256 <KLE>

A:Cross-references: GB:AE001074; GB:AE000782; NID:g2689397; PIDN:AAB90798.1; PID:g265019

C:Superfamily: naphthate synthase; enoyl-CoA hydratase homology

C:Keywords: carbon-oxygen lyase; fatty acid beta-oxidation; hydro-lyase

F:25-175/Domain: enoyl-CoA hydratase homology <ECH>

Query Match

Best Local Similarity 61.8%; Score 34; DB 1; Length 256;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LNAVSTSP.LL 12

Db 210 LNAVSTSP.LL 220

RESULT 15

I48471

Fc gamma (IgG) receptor high affinity - mouse

C:Species: *Mus musculus* (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000

C:Accession: I48471

R:Prins, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarth, P.M.; Wicker, L.S.; G

Science 260, 695-698, 1993

A:Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for

A:Reference number: I48471; MID:93242399

A:Accession: I48471

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-336 <RES>

A:Cross-references: EMBL:X70980; NID:g311748; PIDN:CAA50311.1; PID:g311749

C:Superfamily: Fc gamma receptor I; immunoglobulin homology

C:Keywords: immunoglobulin receptor

F:128-180/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 61.8%; Score 34; DB 2; Length 336;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VLNAVSTSP.LL 12

Db 207 VLNAVSTSP.LL 218

Search completed: June 4, 2001, 12:15:38
Job time: 271 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:38 ; Search time 37.56 Seconds
(without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-21

Perfect score: 55
Sequence: 1 VNNSVTSPLLE 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	55	100.0	374 1	FCG1_HUMAN
2	38	69.1	347 1	ADH3_SUILO
3	36	65.5	533 1	IMA5_MOUSE
4	35	63.6	240 1	FARR_ECOLI
5	35	63.6	347 1	ADH1_SUILO
6	35	63.6	379 1	GRHR_CLAGA
7	35	63.6	422 1	YF21_METJA
8	35	63.6	536 1	IMA5_HUMAN
9	35	63.6	761 1	EEG1_YEAST
10	34	61.8	367 1	FTSZ_AQUAE
11	34	61.8	404 1	FCG1_MOUSE
12	34	61.8	411 1	CRF1_YARLI
13	34	61.8	538 1	GLB1_YARLI
14	34	61.8	580 1	GBA4_USTMA
15	34	61.8	1121 1	ALP1_SCHPO
16	34	61.8	4036 1	RRPL_DUGBV
17	33	60.0	124 1	Y14K_SSV1
18	33	60.0	237 1	PSA5_ORYSA
19	33	60.0	350 1	VGL1_PVRI
20	33	60.0	438 1	FTSZ_ANASP
21	33	60.0	420 1	FTSZ_SYNS3
22	33	60.0	457 1	FTSZ_PORGI
23	33	60.0	538 1	IMA1_MOUSE
24	33	60.0	986 1	BMP1_HUMAN
25	33	60.0	991 1	BMP1_MOUSE
26	33	60.0	1147 1	NRDC_HUMAN
27	33	60.0	1733 1	RPB1_YEAST
28	33	60.0	1752 1	RPB1_SCHPO
29	33	60.0	1840 1	RPB1_ARATH
30	33	60.0	2452 1	RPB1_PLAUF
31	32.5	59.1	403 1	YA24_METJA
32	32	58.2	211 1	YTSF_SPICI
33	32	58.2	233 1	HEXB_ADE12

ALIGNMENTS

RESULT ID	FCG1_HUMAN	STANDARD	PRT	374 AA.
34	32	58.2	237 1	PS51_ARATH
35	32	58.2	237 1	PS52_ARATH
36	32	58.2	237 1	PSA5_SOTBN
37	32	58.2	259 1	ATP6_YEAST
38	32	58.2	380 1	COAT_CNV
39	32	58.2	428 1	FXB2_MOUSE
40	32	58.2	442 1	ARDE_XENLA
41	32	58.2	478 1	D2D1_XENLA
42	32	58.2	485 1	YCL1_KLEPN
43	32	58.2	574 1	YGLF_HRSV1
44	32	58.2	609 1	YKDE_CAEEL
45	32	58.2	734 1	GLCB_AGRTU

RESULT 1
FCG1_HUMAN STANDARD; PRT: 374 AA.
AC P12314; P12315;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA
DE RI) (FCRI) (IGG FC RECEPTOR I) (CD64).
GN FCGRIA OR FCGRI OR FCGI OR IGFRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89098339; PubMed=2974947;
RX Allen J.M., Seed B.;
RT "Nucleotide sequence of three cDNAs for the human high affinity Fc
RT receptor (FCRI)."
RL Nucleic Acids Res. 16:11824-11824(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=89100284; PubMed=2911749;
RX Allen J.M., Seed B.;
RT "Isolation and expression of functional high-affinity Fc receptor
RT complementary DNAs."
RL Science 243:378-381(1989).
CC - FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH
CC AFFINITY RECEPTOR.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.
CC - SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC - DATABASE: NAME=PROW; NOTE=CD guide CD64 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd64.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X14356; CAA32537.1;
CC EMBL: X14355; CAA32536.1;
CC PIR: S03018; S03018;
CC PIR: S03019; S03019;
CC PIR: A41357; A41357;
CC PIR: B41357; B41357;
CC HSSP: P12319; 1ALT.
CC MIM: 146760;
CC InterPro: IPR003006;
CC Pfam: PF00047; 1g; 3.

```

CC      EMBL; Z47543; CAAB7591.1; -.
CC      DR HSSP; P00325; IHDX.
CC      DR InterPro; IPR002085; -.
CC      DR InterPro; IPR002328; -.
CC      DR Pfam; PF00107; adh_zinc; 1.
CC      DR PROSITE; PS00059; ADH_ZINC; 1.
CC      KM Oxidoreductase; Zinc; NAD; Methylation.
CC      FT MOD_RS 11 11 METHYLATION (BY SIMILARITY).
CC      FT MOD_RS 213 213 METHYLATION (BY SIMILARITY).
CC      FT METAL 38 38 ZINC (CATALYTIC) (BY SIMILARITY).
CC      FT METAL 68 68 ZINC (CATALYTIC) (BY SIMILARITY).
CC      FT METAL 98 98 ZINC (SECOND ATOM) (BY SIMILARITY).
CC      FT METAL 101 101 ZINC (SECOND ATOM) (BY SIMILARITY).
CC      FT METAL 104 104 ZINC (SECOND ATOM) (BY SIMILARITY).
CC      FT METAL 112 112 ZINC (SECOND ATOM) (BY SIMILARITY).
CC      FT METAL 154 154 ZINC (CATALYTIC) (BY SIMILARITY).
CC      SQ SEQUENCE 347 AA; 37583 MW; 5D3C003792DE60D CRC64;
CC
CC      Query Match 69.1%; Score 38; DB 1; Length 347;
CC      Best Local Similarity 66.7%; Pred. No. 4;
CC      Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
CC
CC      QY 1 VNASVTSPLIE 12
CC      Db 220 VNASVDPPLAE 231
CC
CC      RESULT 3
CC      IMAS_MOUSE
CC      ID IMAS_MOUSE STANDARD; PRT; 533 AA.
CC      AC O35345;
CC      DT 15-JUL-1998 (Rel. 36, Created)
CC      DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC      DT 01-OCT-2000 (Rel. 40, Last annotation update)
CC      DE IMPORTIN ALPHA-6 SUBUNIT (KARYOPHERIN ALPHA-5 SUBUNIT) (IMPORTIN ALPHA
CC      S2).
CC      GN KPNA5 OR KPNA6.
CC      OS Mus musculus (Mouse).
CC      OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC      RN NCBI_Taxid=10090;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RX MEDLINE=98034160; Pubmed=9369227;
CC      RA Tsuji I., Takumi T., Imamoto N., Yoneda Y.;
CC      RT "Identification of novel homologues of mouse importin alpha, the
CC      RT alpha subunit of the nuclear pore-targeting complex, and their
CC      RT tissue-specific expression.";
CC      RT FEBS Lett. 416:30-34(1997).
CC      RL
CC      CC -1- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
CC      CC EITHER A SINGLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
CC      CC SUBSTRATES TO THE NUCLEAR PORE COMPLEX (NPC). THE COMPLEX IS
CC      CC SUBSEQUENTLY TRANSLOCATED THROUGH THE PORE BY AN ENERGY REQUIRING,
CC      CC RAN-DEPENDENT MECHANISM. AT THE NUCLEOPLASMIC SIDE OF THE NPC, THE
CC      CC THREE COMPONENTS SEPARATE AND IMPORTIN-ALPHA AND -BETA ARE RE-
CC      CC EXPORTED FROM THE NUCLEUS TO THE CYTOPLASM.
CC      CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC      CC -1- TISSUE SPECIFICITY: ONLY SLIGHTLY DETECTED IN EHRlich ASCITES
CC      CC TUMOR CELLS. THYMUS AND SKELETAL MUSCLE, CLEARLY DETECTED IN
CC      CC KIDNEY, SPLEEN, LIVER, HEART, AND LUNG. HIGH EXPRESSION IN TESTIS.
CC      CC -1- DOMAIN: CONSISTS OF AN N-TERMINAL HYDROPHILIC REGION, A
CC      CC HYDROPHOBIC CENTRAL REGION COMPOSED OF 8 REPEATS, AND A SHORT
CC      CC HYDROPHILIC C-TERMINUS. THE N-TERMINAL HYDROPHILIC REGION CONTAINS
CC      CC THE IMPORTIN BETA BINDING DOMAIN (1BB DOMAIN), WHICH IS SUFFICIENT
CC      CC FOR BINDING IMPORTIN BETA AND ESSENTIAL FOR NUCLEAR PROTEIN
CC      CC IMPORT.
CC      CC -1- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
CC      CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


GN M01521.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bolt C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overhaug R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE NA(+/H(+)) EXCHANGER FAMILY. STRONG, TO
CC M.JANNSCHII M0057.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U6593; AAB9540.1;
CC TIGR: M01521;
CC InterPro: IPR000676;
CC Pfam: PF00999; Na_H_Exchange; 1.
CC Hypothetical protein; Transmembrane; Sodium transport; Transport;
KW Symport.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 183 203 POTENTIAL.
FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT TRANSMEM 341 361 POTENTIAL.
FT TRANSMEM 384 404 POTENTIAL.
SQ SEQUENCE 422 AA; 45809 MW; AF56570C626D339 CRC64;

Query Match 63.6%; Score 35; DB 1; Length 422;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASTVSPLE 12
Db 292 VLNASTVSPLE 303

RESULT 8
ID IMAS_HUMAN STANDARD; PRT; 536 AA.
AC 015131;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE IMPORTIN ALPHA-6 SUBUNIT (KARYOPHERIN ALPHA-5 SUBUNIT).
GN KRNAS.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98055463; PubMed=9395085;
RA Koehler M., Ansseau S., Prehn S., Leutz A., Haller H., Hartmann E.;
RT "Cloning of two novel human importin-alpha subunits and analysis of
RT the expression pattern of the importin-alpha protein family.";
RL FEBS Lett. 417:104-108(1997).
CC -1- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
CC SUBSTRATES TO THE NUCLEAR PORE COMPLEX (NPC). THE COMPLEX IS
CC SUBSEQUENTLY TRANSLOCATED THROUGH THE PORE BY AN ENERGY REQUIRING,
CC RAN-DEPENDENT MECHANISM. AT THE NUCLEOPLASMIC SIDE OF THE NPC, THE
CC THREE COMPONENTS SEPARATE AND IMPORTIN-ALPHA AND -BETA ARE RE-
CC EXPORTED FROM THE NUCLEUS TO THE CYTOPLASM.
CC -1- SUBUNIT: FORMS A COMPLEX WITH IMPORTIN BETA-1 SUBUNIT.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DOMAIN: CONSISTS OF AN N-TERMINAL HYDROPHILIC REGION, A
CC HYDROPHOBIC CENTRAL REGION COMPOSED OF 8 REPEATS, AND A SHORT
CC HYDROPHILIC C-TERMINUS. THE N-TERMINAL HYDROPHILIC REGION CONTAINS
CC THE IMPORTIN BETA BINDING DOMAIN (IBB DOMAIN), WHICH IS SUFFICIENT
CC FOR BINDING IMPORTIN BETA AND ESSENTIAL FOR NUCLEAR PROTEIN
CC IMPORT.
CC -1- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF005361; AAC51868.1;
CC MIM: 604545;
CC InterPro: IPR000225;
CC InterPro: IPR002652;
CC Pfam: PF00514; Armadillo_seg; 8.
CC Pfam: PF01749; IBB; 1.
CC PROSITE: PSS0176; ARM_REPEAT; 2.
KW Transport; Protein transport; Repeat.
FT DOMAIN 10 50 IBB.
FT REPEAT 116 159 ARM 1.
FT REPEAT 160 204 ARM 2.
FT REPEAT 205 243 ARM 3.
FT REPEAT 244 288 ARM 4.
FT REPEAT 289 328 ARM 5.
FT REPEAT 329 370 ARM 6.
FT REPEAT 371 410 ARM 7.
FT REPEAT 411 453 ARM 8.
FT DOMAIN 25 28 POLY-ARG.
FT DOMAIN 454 536 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 536 AA; 60266 MW; 4E80B1FAD2A420E CRC64;

Query Match 63.6%; Score 35; DB 1; Length 536;
Best Local Similarity 58.3%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 VLNASTVSPLE 12
Db 205 VLNASTVSPLE 216

RESULT 9
ID EFG1_YEAST STANDARD; PRT; 761 AA.
AC P25039; O99360;
DT 01-MAY-1992 (Rel. 22, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE ELONGATION FACTOR G 1, MITOCHONDRIAL PRECURSOR (MEF-G-1).
 GN MEF1 OR YL069C OR L2195.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OX Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92037620; PubMed=1935960;
 RA Vambutas A., Ackerman S.H., Tzagoloff A.;
 RT "Mitochondrial translational initiation and elongation factors in
 RT Saccharomyces cerevisiae."
 RL Eur. J. Biochem. 201:643-652(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Andre B., Urrestarazu L.A.;
 RT Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X58378; CAA1267.1; -
 DR EMBL: X94607; CAA64315.1; -
 DR EMBL: Z73241; CAA97626.1; -
 DR PIR: S17025; S17025.
 DR PIR: S20179; S20179.
 DR HSSP: P13551; 1E1O.
 DR SGD: S0004059; MEF1.
 DR InterPro: IPR000640; -
 DR InterPro: IPR000795; -
 DR Pfam: PF00679; EFG_C.1.
 DR Pfam: PF00009; GTP_EFPU.1.
 DR PRINTS: PR00315; ELONGATNFCR.
 DR PROSITE: PS00301; ERFACOR_GTP.1.
 DR Elongation factor; Protein biosynthesis; Mitochondrion;
 KW Elongation factor; Protein biosynthesis; Mitochondrion;
 KW Transist peptide; GTP-binding.
 FT CHAIN 1 761
 FT NP_BIND 77 761
 FT NP_BIND 148 84
 FT NP_BIND 202 152
 FT NP_BIND 202 205
 FT CONFLICT 66 66
 FT CONFLICT 233 233
 FT CONFLICT 478 478
 FT CONFLICT 629 629
 FT SEQUENCE 761 AA; 84573 MW; 3E2C534509B09103 CRC64;
 SQ
 Query Match 63.6%; Score 35; DB 1; Length 761;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

AC 066809;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CELL DIVISION PROTEIN FTSZ.
 GN FTSZ OR AO_525.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=VF5;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus."
 RL Nature 392:353-358(1998).
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
 CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
 CC THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
 CC AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
 CC BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
 CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
 CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000694; AAC06771.1; -
 DR InterPro: IPR000158; -
 DR InterPro: IPR003008; -
 DR Pfam: PF00091; tubulin.1.
 DR PRINTS: PR00423; CELLDIVISFTSZ.
 DR PROSITE: PS01134; FTSZ_1; 1.
 DR PROSITE: PS01135; FTSZ_2; 1.
 KW Cell division; Septation; GTP-binding.
 FT NP_BIND 100 108
 FT SEQUENCE 367 AA; 40161 MW; 0DEC8A6A865348A9 CRC64;
 SQ
 Query Match 61.8%; Score 34; DB 1; Length 367;
 Best Local Similarity 87.5%; Pred. No. 27;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RP SEQUENCE FROM N.A.
 RX MEDLINE=90111035; PubMed=2136886;
 RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RT "Molecular cloning and expression of the mouse high affinity Fc
 RT receptor for IgG.";
 RL J. Immunol. 144:371-378(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92163399; PubMed=1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RT "Structure and mapping of the gene encoding mouse high affinity Fc
 RT gamma RI and chromosomal location of the human Fc gamma RI gene.";
 RL J. Immunol. 148:1570-1575(1992).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN GAMMA. HIGH
 CC AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M31314; AAA40056.1; -
 CC PIR: A43511; A43511.
 CC PIR: A46480; A46480.
 CC HSSP: P12319; 1ALT.
 CC MGD: MGI:93498; Fcgr1.
 CC DR InterPro: IPR003006; -
 CC DR Pfam: PF00047; 1g; 3.
 CC KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 CC Immunoglobulin domain.
 CC FT SIGNAL 1 24
 CC FT CHAIN 25 404
 CC FT DOMAIN 25 297 HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC
 CC FT TRANSMEM 298 320 RECEPTOR I.
 CC FT DOMAIN 321 404 EXTRACELLULAR (POTENTIAL).
 CC FT DOMAIN 46 66 POTENTIAL.
 CC FT DOMAIN 67 124 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 154 216 IG-LIKE C2-TYPE DOMAIN (TRUNCATED).
 CC FT CARBOHYD 28 216 IG-LIKE C2-TYPE DOMAIN.
 CC FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 69 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 404 AA: 44887 MW: 1C4F003842767E7 CRC64;
 CC SQ SEQUENCE

Query Match 61.8%; Score 34; DB 1; Length 404;
 Best Local Similarity 66.7%; Pred. No. 30;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VLNAVSTSPLE 12
 DB 202 VLNAVSSPPE 213

RESULT 12
 CRFL YARLI STANDARD; PRT; 411 AA.
 AC P45815;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE COPPER RESISTANCE PROTEIN CRFL.
 GN CRFL.
 OS Yarrowia lipolytica (Candida lipolytica).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OC NCBI_TaxID=4952;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 20460 / W29;
 RA Prado-Gonzalez M.;
 RL Thesis (1994), University of Salamanca, Spain.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS A "COPPER-FIST" DNA-BINDING DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z23265; CA80803.1; -
 CC DR InterPro: IPR001083; -
 CC DR Pfam: PF00649; Copper-fist; 1.
 CC DR PRINTS: PR00617; COPPERFIST.
 CC DR PROSITE: PS01119; COPPER_FIST_1; 1.
 CC DR PROSITE: PS50073; COPPER_FIST_2; 1.
 CC KW Transcription regulation; DNA-binding; Copper; Nuclear protein.
 CC FT DNA_BIND 1 40
 CC FT DOMAIN 63 66 COPPER-FIST.
 CC FT DOMAIN 126 140 POLY-SER.
 CC FT DOMAIN 205 211 POLY-GLN.
 CC FT DOMAIN 205 211 POLY-SER.
 CC SQ SEQUENCE 411 AA: 43688 MW: 9AD602FA467EAC7 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 411;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNAVSTSPLE 12
 DB 272 ILNDLSSPLD 283

RESULT 13
 ID GLE1_YEAST STANDARD; PRT; 538 AA.
 AC Q12315;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE RNA EXPORT FACTOR GLE1.
 GN GLE1 OR BRR3 OR YDL207W OR D1049.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RX [1]
 RP SEQUENCE FROM N.A. AND MUTAGENESIS.
 RC STRAIN=S288C / FY1679;
 RA Bahr A., Moeller-Rieker S., Hankeln T., Kraemer C., Schmidt E.R.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RNA EXPORT FACTOR. THE MECHANISM MAY INVOLVE THE DIRECT
 CC TARGETING OF RNA/RNP MOLECULES TO OR THROUGH THE NUCLEAR PORE
 CC COMPLEX BY MEANS OF RPL1 AND NUP100 INTERACTIONS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. LOCALIZED PREDOMINANTLY AT THE
 CC NUCLEAR PORE COMPLEXES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U68475; AAC49444.1; -
 DR EMBL: X99000; CAA67484.1; -
 DR EMBL: Z74255; CAA8785.1; -
 DR SGD: S0002366; GLE1.
 KW Nuclear protein; mRNA processing.
 FT MUTAGEN 351 358 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT MUTAGEN 353 353 L->A: PARTIAL LOSS OF ACTIVITY.
 FT MUTAGEN 356 353 L->A: PARTIAL LOSS OF ACTIVITY.
 FT MUTAGEN 358 356 L->A: TEMPERATURE SENSITIVE.
 FT MUTAGEN 358 358 L->A: PARTIAL LOSS OF ACTIVITY.
 SQ SEQUENCE 538 AA; 62073 MW; C9C6B1513AF95711 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 538;
 Best Local Similarity 60.0%; Pred. No. 42;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 NASTVSPLE 12
 DB 116 NSMATAPILE 125

RESULT 14
 GB44_USTMA STANDARD; PRT; 580 AA.
 AC P87035:

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN ALPHA-4 SUBUNIT.
 GN GP44.

OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_Taxid=5270;

RP SEQUENCE FROM N.A.
 RN STRAIN=FBI;

RA MEDLINE=97299870; PubMed=9155019;
 RA Regenfelder E., Speilg T., Hartmann A., Lauenstein S., Bolker M.,
 RA Kahmann R.;
 RA "G proteins in Ustilago maydis: transmission of multiple signals?";
 RL EMBL J. 16:1934-1942(1997).

CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.

CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.

CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U68778; AAC49727.1; -
 DR HSSP: P10824; IAS3.
 DR InterPro: IPR001019; -
 DR Pfam: PF00503; G-alpha.1.
 DR PRINTS: PR00318; GPROTEIN.
 KW GTP-binding; Transducer; Multigene family.
 FT NP_BIND 101 108 GTP (BY SIMILARITY).
 FT NP_BIND 415 419 GTP (BY SIMILARITY).

FT NP_BIND 484 487 GTP (BY SIMILARITY).
 SQ SEQUENCE 580 AA; 65198 MW; 7EA38FBC0FACFCC CRC64;

Query Match 61.8%; Score 34; DB 1; Length 580;
 Best Local Similarity 58.3%; Pred. No. 46;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VNASTVSPLE 12
 DB 464 LFMQIVTNPILE 475

RESULT 15
 ALP1_SCHPO STANDARD; PRT; 1121 AA.

AC 010197; P78749;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN-FOLDING COFACTOR D.

GN ALP1 OR SPBC11C11.04C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OX NCBI_Taxid=4896;

RP SEQUENCE FROM N.A.
 RN STRAIN=972 / HM123;
 RC MEDLINE=98119809; PubMed=9450991;
 RA Hirata D., Masuda H., Edlison M., Toda T.;
 RT "Essential role of tubulin-folding cofactor D in microtubule assembly
 RL and its association with microtubules in fission yeast.";
 RL EMBL J. 17:658-666(1998).

RP SEQUENCE FROM N.A.
 RN STRAIN=972;

RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: HAS A FUNCTION IN THE FOLDING OF BETA-TUBULIN; MAY PLAY
 CC A VITAL ROLE IN MICROTUBULE-DEPENDENT PROCESSES AS A MICROTUBULE-
 CC ASSOCIATED PROTEIN.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: Y10106; CAA71193.1; -
 DR EMBL: AL031528; CAA20686.1; -
 KW Microtubules; Transmembrane; Glycoprotein.

FT TRANSMEM 129 149 POTENTIAL.
 FT TRANSMEM 308 328 POTENTIAL.
 FT TRANSMEM 378 398 POTENTIAL.
 FT TRANSMEM 581 601 POTENTIAL.
 FT TRANSMEM 653 673 POTENTIAL.
 FT TRANSMEM 896 916 POTENTIAL.
 FT TRANSMEM 962 982 POTENTIAL.
 FT CARBOHYD 122 122 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 373 373 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 721 721 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 885 885 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1099 1099 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 988 1001 MISSING (IN REF. 2).
 SQ SEQUENCE 1121 AA; 128186 MW; DCF045EB3B81213A2 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 1121;
Best Local Similarity 60.0%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 LNASTSPPL 11
:| | :|||:
Db 372 VNITSPLV 381

Search completed: June 4, 2001, 12:24:01
Job time: 563 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:37 ; Search time 107.68 Seconds

(without alignments)
13.062 Million cell updates/sec

Title: US-09-284-107-21

Perfect score: 55

Sequence: 1 VLNASTSPLE 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organelle:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-unclassified:*
13: SP-vertebrate:*
14: SP-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	374	4	Q92663 homo sapien
2	55	100.0	375	4	Q92495 homo sapien
3	40	72.7	1199	14	Q83859 nllaparvata
4	38	69.1	321	6	P79119 bos taurus
5	38	69.1	322	4	Q99645 homo sapien
6	37	67.3	175	2	O53805 mycobacteri
7	37	67.3	361	1	O27116 methanobact
8	37	67.3	1118	8	O21376 pichia kluy
9	37	67.3	1424	10	Q9L7R5 arbidopsi
10	36	65.5	52	5	Q9NMU0 leishmania
11	36	65.5	136	5	Q9NMU8 leishmania
12	36	65.5	536	4	O60684 homo sapien
13	36	65.5	566	4	Q9NU85 homo sapien
14	35	63.6	44	14	P87859 hepatitis c
15	35	63.6	44	14	P87859 hepatitis c
16	35	63.6	366	14	Q9QBC3 yaba monkey
17	35	63.6	536	4	Q9NMU0 homo sapien
18	35	63.6	591	10	Q9LYA6 arabidopsi
19	35	63.6	632	2	O83428 treponema p

20	35	63.6	662	14	P89940 citrus mosa
21	35	63.6	922	14	Q9JGK6 citrus mosa
22	35	63.6	1223	3	Q02979 saeccharomy
23	35	63.6	1455	5	O17012 caenorhabd
24	35	63.6	3010	14	Q9QP06 hepatitis c
25	35	63.6	4124	5	O97218 leishmania
26	34	61.8	129	1	Q9UXA1 sulfolobus
27	34	61.8	226	2	Q9RLU5 lactococcus
28	34	61.8	256	1	O29814 archaeglob
29	34	61.8	387	11	Q9JMA2 mus musculu
30	34	61.8	393	2	O85785 synechococ
31	34	61.8	402	14	Q98233 molluscum c
32	34	61.8	423	5	Q9TE37 caenorhabd
33	34	61.8	570	2	O51775 borrelia bu
34	34	61.8	581	5	O9VWE5 drosophila
35	34	61.8	803	5	O9Y163 drosophila
36	34	61.8	829	10	Q9LFN5 arbidopsi
37	34	61.8	1011	5	Q9Y148 drosophila
38	34	61.8	1059	5	Q9V3X6 drosophila
39	33	60.0	76	5	O02324 caenorhabd
40	33	60.0	134	2	O50135 ruminococcu
41	33	60.0	145	2	O34084 streptococ
42	33	60.0	145	9	Q9MBX8 streptococ
43	33	60.0	151	5	Q9W436 drosophila
44	33	60.0	156	1	Q9YAB4 aeropyrum p
45	33	60.0	201	10	Q9LY44 arabidopsi

ALIGNMENTS

RESULT 1
Q92663 PRELIMINARY; PRT; 374 AA.
AC Q92663;
ID 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE FC GAMMA RECEPTOR I.
GN A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93055454; PubMed=1430234;
RA Porjes A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RT Kimberly R.P.,
RT "Novel Fc gamma receptor I family gene products in human mononuclear
cells".
RL J. Clin. Invest. 90:2102-2109(1992).
DR EMBL: L03418; AAA36049.1.
DR HSSP: P12319; IALT.
DR INTERPRO: IPR003006;
DR PFAM: PF00047; Ig_3.
DR PRODOM: PD002534; 1.
SQ SEQUENCE 374 AA; 42632 MW; D33D59398CEEA699 CRC64;

Query Match 100.0%; Score 55; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0094;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLNASTSPLE 12
Db 193 VLNASTSPLE 204
RESULT 2
Q92495 PRELIMINARY; PRT; 375 AA.

AC Q92495;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE FC GAMMA RECEPTOR TYPE I (FC GAMMA RIB-FC GAMMA RECEPTOR).
 GN CD64 OR FC-GAMMA-RIB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RA Benesch P.D., Sasstry K.N., Iyer R.R., Eichbaum Q.G., Raveh D.,
 RA Ezekowitz A.B.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-224 FROM N.A.
 RX MEDLINE=93018827; PubMed=1402657;
 RA Benesch P.D., Sasstry K., Iyer R.R., Eichbaum Q.G., Raveh D.P.,
 RA Ezekowitz R.A.;
 RL "Definition of interferon gamma-response elements in a novel human FC
 gamma receptor gene (Fc gamma RIB) and characterization of the gene
 structure.";
 RT J. Exp. Med. 176:1115-1123(1992);
 RL DR EMBL: M91555; AA58414.1; JOINED.
 DR EMBL: M91550; AA58414.1; JOINED.
 DR EMBL: M91551; AA58414.1; JOINED.
 DR EMBL: M91552; AA58414.1; JOINED.
 DR EMBL: M91553; AA58414.1; JOINED.
 DR EMBL: M91554; AA58414.1; JOINED.
 DR EMBL: M91555; AA58414.1; JOINED.
 DR EMBL: S45709; AAD13842.1; JOINED.
 DR EMBL: S45707; AAD13842.1; JOINED.
 DR EMBL: S45708; AAD13842.1; JOINED.
 DR EMBL: S45704; AAD13842.1; JOINED.
 DR EMBL: S45705; AAD13842.1; JOINED.
 DR HSSP: P12319; 1ALT.
 DR INTERPRO: IPR003006;
 DR PFM: PF00047; 1g: 3;
 DR PRODOM: PD002534; 1;
 SQ SEQUENCE 375 AA; 42861 MW; A84D464C70DD0F91 CRC64;

Query Match 100.0%; Score 55; DB 4; Length 375;
 Best Local Similarity 100.0%; Pred. No. 0.0094;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VINASTVSPLE 12
 DB 194 VINASTVSPLE 205

RESULT 3
 O83859 PRELIMINARY; PRT: 1199 AA.
 AC Q83859;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-JAN-1999 (TREMBLrel. 09, last annotation update)
 DE 136.6 KDA PROTEIN.
 OS Nilaparvata lugens reovirus.
 OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
 OX NCBI_TaxID=33724;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IZUMO;
 RA Nakashima N., Noda H.;
 RT "Nucleotide and deduced amino acid sequences of Nilaparvata lugens
 reovirus: a putative member of the genus Fibivirus.";
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-7 FROM N.A.
 RC STRAIN=IZUMO;

RX MEDLINE=96145144; PubMed=8558122;
 RA Nakashima N., Koizumi M., Matanabe H., Hiroaki N.;
 RT "Complete nucleotide sequence of the Nilaparvata lugens reovirus: a
 putative member of the genus Fibivirus.";
 RL J. Gen. Virol. 77:139-146(1996).
 DR EMBL: D49694; BAA08543.1;
 SQ SEQUENCE 1199 AA; 136614 MW; C581EC33B8AC765A CRC64;

Query Match 72.7%; Score 40; DB 14; Length 1199;
 Best Local Similarity 66.7%; Pred. No. 27;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VINASTVSPLE 12
 DB 753 VINASTVSPLE 764

RESULT 4
 P79119 PRELIMINARY; PRT: 321 AA.
 AC P79119;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE SMALL CHONDROITIN/DERMATAN SULFATE PROTEOGLYCAN PRECURSOR (PG-LB)
 (PGLB) (EPITHELIUM).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RC TISSUE=FETAL EPITHELIUM; CARTILAGE;
 RX MEDLINE=97373567; PubMed=9228042;
 RA Johnson H.J., Rosenberg L., Choi H-U., Garza S., Hook M., Neame P.D.;
 RT "Characterization of epiphysean, a small proteoglycan with a leucine-
 rich repeat core protein.";
 RL J. Biol. Chem. 272:18709-18717(1997).
 CC -1- FUNCTION: MAY HAVE A ROLE IN BONE FORMATION, AND ALSO IN
 ESTABLISHING THE ORDERED STRUCTURE OF CARTILAGE THROUGH MATRIX
 ORGANIZATION.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE ZONE OF
 FLATTENED CHONDROCYTES OF THE DEVELOPING LIMB CARTILAGE.
 CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- PTM: A LONG AND A SHORT FORM; PRESENT IN APPROXIMATELY EQUIMOLAR
 AMOUNTS; MAY ARISE BY PROTEOLYSIS OR CLEAVAGE BY EXOPEPTIDASES.
 CC -1- PTM: THE THREE GLYCOSAMINOGLYCAN CHAINS ARE DERMATAN SULFATE
 (POTENTIAL).
 CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
 MANY PROTEINS. NUMBER IN THIS PROTEIN: 6.
 CC -1- SIMILARITY: IN THE C-TERMINAL TO HUMAN OIF.
 DR EMBL: U77127; AAB68397.1;
 DR INTERPRO: IPR003722;
 DR INTERPRO: IPR001611;
 DR PFM: PF00560; LRR: 3;
 DR PFM: PF01462; LRRNT: 1;
 DR KW Proteoglycan; Glycoprotein; Signal; Repeat; Leucine-repeat;
 KM Connective tissue; Extracellular matrix.
 FT SIGNAL 1 19
 FT PROPEP 20 30
 FT CHAIN 31 321
 FT DOMAIN 152 300
 FT REPEAT 152 166
 FT REPEAT 167 190
 FT REPEAT 191 210
 FT REPEAT 211 236
 FT REPEAT 237 257
 FT REPEAT 258 288
 FT REPEAT 289 300
 FT SMALL CHONDROITIN/DERMATAN SULFATE
 PROTEOGLYCAN.
 FT 7 LEUCINE-RICH TANDEM REPEATS.
 FT LRR 1 (INCOMPLETE).
 FT LRR 2.
 FT LRR 3.
 FT LRR 4.
 FT LRR 5.
 FT LRR 6.
 FT LRR 7 (INCOMPLETE).

FT DISULFID 117 129 BY SIMILARITY.
 FT DISULFID 278 311
 FT DOMAIN 77 86 POLY-GLU.
 FT CARBOHYD 60 60 GLCNAC (POTENTIAL).
 FT CARBOHYD 64 64 GLCNAC (POTENTIAL).
 FT CARBOHYD 95 95 GLCNAC (POTENTIAL).
 FT CARBOHYD 282 282 PROBABLE.
 FT CARBOHYD 319 319 GLCNAC (POTENTIAL).
 SQ SEQUENCE 321 AA: 36687 MW: 5D558F31CDB1F89 CRC64:

Query Match 69.1%; Score 38; DB 6; Length 321;
 Best Local Similarity 58.3%; Pred. No. 18;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLNASTSPLE 12
 DB 13 ILDAVATAPLE 24

RESULT 5
 ID 099645 PRELIMINARY; PRT: 322 AA.
 AC 099645;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE SMALL CHONDROITIN/DERMATAN SULFATE PROTEOGLYCAN-PRECURSOR (PG-LB)
 DE (PGLB) (EPHRYCAN) (DERMATAN SULFATE PROTEOGLYCAN 3) (DSPG3).
 GN DSPG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC TISSUE-CHONDROCYTES;
 RX MEDLINE-97131519; PubMed-8975717;
 RA Deere M., Johnson J., Garza S., Harrison W.R., Yoon S.-J.,
 RA Elder F.F.B., Kucherlapati R., Hook M., Hecht J.T.;
 RT Characterization of human DSPG3, a small dermatan sulfate
 RT proteoglycan.";
 RL Genomics 38:399-404(1996).
 CC -1- FUNCTION: MAY HAVE A ROLE IN BONE FORMATION. AND ALSO IN
 CC ESTABLISHING THE ORDERED STRUCTURE OF CARTILAGE THROUGH MATRIX
 CC ORGANIZATION.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: CARTILAGE, LIGAMENT, AND PLACENTA.
 CC -1- PTM: THE THREE GLYCOSAMINOGLYCAN CHAINS ARE DERMATAN SULFATE
 CC (POTENTIAL).
 CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
 CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 6.
 CC -1- SIMILARITY: TO OTHER SMALL CHONDROITIN/DERMATAN SULFATE
 CC PROTEOGLYCANS.
 DR EMBL: U59111; AAC50945.1;
 DR MIM: 601657;
 DR INTERPRO: IPR000372;
 DR INTERPRO: IPR001611;
 DR PFAM: PF00560; LRR: 3;
 DR PFAM: PF01462; LRRT: 1;
 KW Proteoglycan; Glycoprotein; Signal; Repeat; Leucine-repeat;
 KW Connective tissue; Extracellular matrix.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 30 POTENTIAL.
 FT CHAIN 31 322 SMALL CHONDROITIN/DERMATAN SULFATE
 FT PROTEOGLYCAN.
 FT DOMAIN 153 301 LRR 1 (INCOMPLETE).
 FT REPEAT 153 167 7 LEUCINE-RICH TANDEM REPEATS.
 FT REPEAT 168 191 LRR 2.
 FT REPEAT 192 211 LRR 3.
 FT REPEAT 212 237 LRR 4.
 FT REPEAT 238 258 LRR 5.
 FT REPEAT 259 289 LRR 6.

FT REPEAT 290 301 LRR 7 (INCOMPLETE).
 FT DISULFID 118 130 BY SIMILARITY.
 FT CARBOHYD 279 312 GLCNAC (POTENTIAL).
 FT CARBOHYD 64 64 POLY-GLU.
 FT DOMAIN 81 87 GLCNAC (POTENTIAL).
 FT CARBOHYD 96 96 GLCNAC (POTENTIAL).
 FT CARBOHYD 283 283 POTENTIAL.
 FT CARBOHYD 302 302 POTENTIAL.
 FT CARBOHYD 320 320 GLCNAC (POTENTIAL).
 SQ SEQUENCE 322 AA: 36607 MW: E42E7C4F2/C858A CRC64:

Query Match 69.1%; Score 38; DB 4; Length 322;
 Best Local Similarity 58.3%; Pred. No. 18;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLNASTSPLE 12
 DB 13 ILDAVATAPLE 24

RESULT 6
 ID 053805 PRELIMINARY; PRT: 175 AA.
 AC 053805;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE PERS-FAMILY PROTEIN.
 DE RV0742 OR MTV041.16.
 GN Mycobacterium tuberculosis;
 OS Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37R;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL: AL021958; CAI17509.1;
 DR TUBERCULIST; RV0742;
 DR INTERPRO: IPR000084;
 DR PFAM: PF00934; PE: 1.
 DR PRODOM: PD001223; -; 1.
 SQ SEQUENCE 175 AA: 15545 MW: C62849BB8B51AAE9 CRC64;

Query Match 67.3%; Score 37; DB 2; Length 175;
 Best Local Similarity 77.8%; Pred. No. 15;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 ASYSTSPLE 12
 DB 96 ASITSPLE 104

RESULT 7
 ID 027116 PRELIMINARY; PRT: 361 AA.
 AC 027116;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)

DE HYPOTHETICAL 39.9 KDA PROTEIN.
 GN MTH1037.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 OX NCBI_TaxID=2166;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H.
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spatafora R., Viçare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Caruso A., Bush D., Saffer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noilting J., Reeve J.N.,
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT J. Bacteriol. 179:7135-7155(1997)."
 DR EMBL: AE000876; AAB85533.1; -
 DR INTERPRO: IPR000130; -
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR Hypothetical protein.
 KW SEQUENCE 361 AA; 39939 MW; 66C43939592AD14 CRC64;
 SQ

Query Match 67.3%; Score 37; DB 1; Length 361;
 Best Local Similarity 72.7%; Pred. No. 32;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNASVTSPL 11
 ||| ||| |||
 DB 172 VLNKSVLSPML 182

RESULT 8
 ID 021376 PRELIMINARY; PRT: 1118 AA.
 AC 021376;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE DNA POLYMERASE.
 OS Pichia kluyveri (Yeast).
 OC Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=36015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CBS 7907;
 RA Blaisonneau J., Nosek J., Fukuhara H.;
 RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Y11606; CAA72340.1; -
 DR INTERPRO: IPR002064; -
 DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN_1.
 DR Mitochondrion.
 KW SEQUENCE 1118 AA; 130725 MW; 773C5756844E3944 CRC64;
 SQ

Query Match 67.3%; Score 37; DB 8; Length 1118;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 1 VNASV--TSPL 12
 ||||| |||||
 DB 551 VLNASLTKTSPLE 564

RESULT 9
 O9LTR5 PRELIMINARY; PRT: 1424 AA.

AC O9LTR5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE SIMILARITY TO BACTERICIDAL PERMEABILITY-INCREASING PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.,
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 DR EMBL: AB024036; BAB02808.1; -
 DR SEQUENCE 1424 AA; 154737 MW; E952BE4F1BB41C96 CRC64;
 SQ

Query Match 67.3%; Score 37; DB 10; Length 1424;
 Best Local Similarity 88.9%; Pred. No. 1.3e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VNASVTSPL 9
 ||| ||| |||
 DB 483 VLNASVTSPL 491

RESULT 10
 ID 09NM00 PRELIMINARY; PRT: 52 AA.
 AC 09NM00;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PROBABLE ADA2-LIKE PROTEIN (FRAGMENT).
 GN LM26.72.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL160493; CAB97690.1; -
 FT NON_TER
 KW SEQUENCE 52 AA; 5639 MW; 0BD2006CA2F21EC5 CRC64;
 SQ

Query Match 65.5%; Score 36; DB 5; Length 52;
 Best Local Similarity 58.3%; Pred. No. 7;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNASVTSPL 12
 ||| ||| |||
 DB 29 VLGANYTHPLMO 40

RESULT 11
 O9NML8 PRELIMINARY; PRT: 136 AA.
 ID 09NML8;
 AC 09NML8;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE POSSIBLE MICROFUBULE-ASSOCIATED PROTEIN (FRAGMENT).
 GN LM26.74.
 OS Leishmania major.
 CC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN.
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL160493; CAB97692.1;
 FT NON_TER
 SO SEQUENCE 136 AA; 15280 MW; 40D5F80759D361A3 CRC64;

Query Match 65.5%; Score 36; DB 5; Length 136;
 Best Local Similarity 58.3%; Pred. No. 18;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASVTSPL 12
 |||:||||:
 Db 64 VLGNVTHPLMQ 75

RESULT 12
 ID 060684 PRELIMINARY; PRT; 536 AA.

AC 060684;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE IMPORIN ALPHA 7 SUBUNIT.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Christiansen M., Hartmann E.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF060543; AAC15233.1;
 DR HSP; Q02821; IBK5.
 DR INTERPRO: IPR000225;
 DR INTERPRO: IPR002652;
 DR PFM; PF00514; Armadillo_seg. 8.
 DR PFM; PF01749; IB; 1.
 SO SEQUENCE 536 AA; 60029 MW; 9D0E27482B9HDED3 CRC64;

Query Match 65.5%; Score 36; DB 4; Length 536;
 Best Local Similarity 63.6%; Pred. No. 74;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASVTSPL 11
 |||:||||:
 Db 205 VLNCSTLNPPL 215

RESULT 13
 ID 09NU85 PRELIMINARY; PRT; 566 AA.

AC 09NU85;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE D5622L5.1 (IMPORTIN ALPHA 7 (KARYOPHERIN)).
 GN D5622L5.1.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Frankland J.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL049795; CAB75614.1;
 SO SEQUENCE 566 AA; 63004 MW; 48657248DF26782 CRC64;

Query Match 65.5%; Score 36; DB 4; Length 566;
 Best Local Similarity 63.6%; Pred. No. 79;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLNASVTSPL 11
 |||:||||:
 Db 235 VLNCSTLNPPL 245

RESULT 14
 ID P87859 PRELIMINARY; PRT; 44 AA.

AC P87859;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE PARTIAL NS4 GENE (PATIENT 2, 1989, CLONE 4) (FRAGMENT).
 GN NS4.
 OS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 CC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-PATIENT 2, 1989;
 RC MEDLINE-97201623; PubMed-9049409;
 RA Devereux H.L., Brown D., Dushenko G.M., Emery V.C., Lee C.A.;
 RT "Long-term evolution of the 5'UTR and a region of NS4 containing a CTL
 epitope of hepatitis C virus in two haemophilic patients."
 RL J. Gen. Virol. 78:583-590(1997).
 DR EMBL; Z84308; CAB06360.1;
 DR INTERPRO: IPR000745;
 DR PFM; PF01006; HCV_NS4a; 1.
 DR NON_TER 1
 FT NON_TER 44
 SO SEQUENCE 44 AA; 4297 MW; 7B755D858C4A100E CRC64;

Query Match 63.6%; Score 35; DB 14; Length 44;
 Best Local Similarity 77.8%; Pred. No. 9.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LNASVTSPL 10
 |||:||||:
 Db 4 LTASITSP 12

RESULT 15
 ID P87883 PRELIMINARY; PRT; 44 AA.

AC P87883;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE PARTIAL NS4 GENE (PATIENT 1, 1994, CLONE 3) (FRAGMENT).
 GN NS4.
 OS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 CC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-PATIENT 1, 1994;
 RC MEDLINE-97201623; PubMed-9049409;
 RA Devereux H.L., Brown D., Dushenko G.M., Emery V.C., Lee C.A.;
 RT "Long-term evolution of the 5'UTR and a region of NS4 containing a CTL

RT epitope of hepatitis C virus in two haemophilic patients."

RL J. Gen. Virol. 78:583-590(1997).

DR EMBL; Z84361; CAB06413.1;

DR INTERPRO; IPR00745;

DR PFAM; PF01006; HCV_NS4a; 1.

FT NON_TER 1 44 44

SEQUENCE 44 AA; 4317 MW; 71D5F7858B3A670E CRC64;

Query Match

Best Local Similarity 63.6%; Score 35; DB 14; Length 44;
Pred. NO. 9.3;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LNASVTSPL 10
| | | | |
| | | | |
Db 4 LTRASITSPL 12

Search completed: June 4, 2001, 12:23:21
Job time: 584 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:26 ; Search time 106.14 Seconds
(without alignments)
6:463 Million cell updates/sec

Title: US-09-284-107-22

Perfect score: 61

Sequence: 1 GNLVTLSCETKL 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A: Geneseq.0401:*

- 1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	12	19	W60554
2	61	100.0	261	20	W33183
3	61	100.0	344	13	R22549
4	61	100.0	344	17	R91439
5	61	100.0	344	21	Y96183
6	61	100.0	374	13	R20811
7	61	100.0	374	13	R22550
8	61	100.0	374	17	R91438
9	61	100.0	374	17	W00859
10	61	100.0	374	19	W60448
11	61	100.0	374	19	W97833

12	61	100.0	374	19	W97834	Human Fc receptor
13	61	100.0	374	21	Y96134	Human macrophage-s
14	61	100.0	374	21	Y96226	Human high affinity
15	399	61	100.0	21	B43683	Human cancer assoc
16	48	78.7	410	12	R12428	Hybrid Fc(gamma)RI
17	43	70.5	359	20	Y41690	Human PRO329 prote
18	43	70.5	359	21	B44246	Human PRO329 (UNO2
19	43	70.5	377	20	W61195	Human Fc receptor
20	42	68.9	358	21	Y94336	Human cell surface
21	42	68.9	359	21	B34744	Human secreted pro
22	39	63.9	273	19	W98635	H. pylori GHPD 324
23	38	62.3	547	20	Y40905	Arabidopsis thalia
24	38	62.3	584	21	B42539	Human ORFX ORF2303
25	37	60.7	94	16	R80528	Human T cell recep
26	37	60.7	113	13	R26087	Immunising peptide
27	37	60.7	113	13	R26147	Vbeta17 chain of T
28	37	60.7	113	14	R38722	Vbeta17 variable r
29	37	60.7	119	19	W76993	Human T cell recep
30	37	60.7	119	17	R90835	1066 heavy chain v
31	37	60.7	214	20	Y41178	Llama Vhh polypept
32	37	60.7	287	21	Y36056	HLA-A2/Flu restric
33	37	60.7	287	21	Y56077	HLA-A2/Flu restric
34	37	60.7	287	21	Y57853	TCR beta chain and
35	37	60.7	287	21	Y57865	TCR beta chain and
36	37	60.7	305	21	Y56057	HLA-A2/Flu restric
37	37	60.7	305	21	Y57866	TCR beta chain and
38	37	60.7	391	17	R97659	Sling chain T cel
39	37	60.7	434	20	Y13567	Human Robo 2 polyp
40	37	60.7	434	20	Y08405	Human partial ROBO
41	36	59.0	175	19	W53283	Amino acid sequenc
42	36	59.0	194	19	W53282	Amino acid sequenc
43	36	59.0	670	22	B51085	Cabbage looper OA/
44	35	57.4	115	21	B24946	Plant SDF encoded
45	35	57.4	117	21	B24945	Plant SDF encoded

ALIGNMENTS

RESULT 1	
ID W60554	standard; peptide: 12 AA.
XX	
AC W60554:	
XX	
DT 18-AUG-1998 (first entry)	
XX	
DE Oligopeptide from extracellular domain of CD64.	
XX	
KW Extracellular domain; CD64; identification; antibody;	
KW Immunohistochemical; immunofluorescent analysis; detection;	
KW cell transformation; mutation; anti; oncogene.	
XX	
OS Synthetic.	
XX	
PN W09815833-A1.	
XX	
PD 16-APR-1998.	
XX	
PF 07-OCT-1997; 97NO-NL00557.	
XX	
PR 08-OCT-1996; 96EP-0202791.	
XX	
PA (UYUT-) RIKUSUNIV UTRECHT.	
XX	
PI De Krulff CA, Logtenberg T;	
XX	
DR WPI; 1998-240964/21.	
XX	
PT Identifying peptide(s) binding specifically to protein target - by	
PT expressing on phage surface and testing for binding to immobilised	
PT oligopeptide derived from the target, useful for, e.g. identifying	
PT specific antibodies	

XX Example 1; Page 29; 40pp; English.

CC Synthetic oligopeptides W60537-61 are derived from the extracellular
CC domain of CD64. They were synthesised on 25 polyethylene rods as
CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
CC affinity binding to the CD64-covered rods. The specification describes a
CC method for the identification of a peptide able to bind specifically to a
CC target protein. The method comprises displaying the peptide on the
CC surface of a replicable display package, synthesising oligopeptides
CC derived from the target protein on a solid phase, and testing for binding
CC between the peptide and oligopeptides. The method is used to screen large
CC peptide libraries, especially to detect antibodies, or their fragments,
CC that bind to cell markers or that can differentiate between different
CC forms of the same protein, including bispecific antibodies that bind to
CC two non-overlapping epitopes on the same monomeric antigen or two
CC epitopes on different molecules. The genes/oligonucleotides that encode
CC selected peptides can be isolated and used for recombinant production of
CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
CC immunofluorescent analysis, and also to detect cell transformation caused
CC by mutation in (anti)oncogenes.

SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
| | | | | | | | | | | | | |
Db 1 gnlvtlscetkl 12

RESULT 2

Y33183
ID Y33183 standard; Protein: 261 AA.

XX Y33183;

DT 15-NOV-1999 (first entry)

DE Human sFc-gammaRI protein fragment.

XX Fc receptor; Fc-gammaRI; human; FcR; model; three-dimension: 3-D;
XX atomic coordinate; bioactive compound design; computer assisted;
XX drug design; therapy; inhibitor; Fc-gamma; Fc-epsilon; IgG; IgE;
XX tissue damage; hypersensitivity; inflammatory cell recruitment;
XX inflammatory modulator; Fc-gammaRI; immune function regulation;
XX anti-inflammatory; immunoprotective; sFc-gammaRI.

OS Homo sapiens.

PN WO9940117-A1.

PD 12-AUG-1999.

PF 04-FEB-1999; 99MO-IB00367.

PR 11-SEP-1998; 98US-0099994.

PR 06-FEB-1998; 98US-0073972.

PA (ILEX-) ILEXUS PTY LTD.

PI Baeil JB, Epa V, Garrett TPJ, Hogarth PM, Matthews BR;
PI Maxwell KF, McCarthy TD, McKenzie IC, Pietersz GA;
PI Powell MS;

PI WPI; 1999-539978/45.

XX Three-dimensional structures and models of Fc receptors, useful in
XX computer-assisted drug design

PS Claim 11; Page 316-317; 326pp; English.

CC This invention describes a novel model of an Fc receptor (FcR) protein
CC representing a three-dimensional (3-D) structure that substantially
CC conforms to the specified atomic coordinates. Computer model images of
CC the FcR can be used to design bioactive chemical compounds, e.g.
CC oligonucleotides, peptides, peptidomimetics and small organic molecules,
CC by computer-assisted methods of drug design. Therapeutic compositions
CC that inhibit the activity of Fc-gammaR or Fc-epsilonR can be used to
CC reduce IgG- or IgE-mediated tissue damage, e.g. IgG- or IgE-mediated
CC hypersensitivity, recruitment of inflammatory cells or release of
CC inflammatory mediators. The therapeutic compositions can also be used to
CC treat disease in animals. The knowledge of the 3-D structure of Fc-gamma
CC RIa and models of other FcR provides a means for designing and producing
CC compounds that regulate immune function and inflammation in an animal,
CC including humans (i.e. structure based drug design). For example,
CC chemical compounds can be designed to block binding of immunoglobulin to
CC an Fc receptor protein using various computer programs and models. The
CC products of the invention have anti-inflammatory and immunoprotective
CC activity. This sequence represents the human sFc-gammaRI protein
CC fragment described in the method of the invention.

SQ Sequence 261 AA;

Query Match 100.0%; Score 61; DB 20; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.0073;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
| | | | | | | | | | | | | |
Db 186 gnlvtlscetkl 197

RESULT 3

R22549
ID R22549 standard; Protein: 344 AA.

XX R22549;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FcRI receptor encoded by clone p98.

XX Rapid immunoselection cloning technique; cell surface-antigen;
XX immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN WO9201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90MO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GENO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amiot M;

PI WPI; 1992-056864/07.

PI N-PSDB; Q21179.

PI New CD53 cell surface antigen and DNA encoding it - for
PI immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

XX This amino acid sequence was predicted from the cDNA sequence of
XX cDNA clone p98/X2. It differs from the sequence predicted from
XX clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
XX while clone p98/X2 predicts a Thr residue. At position 58, p90

CC (see Q21180) predicts Val and p98/X2 predicts Leu. Sequences
 CC predicted from all 3 clones show the typical features of a type I
 CC integral membrane protein and include a short hydrophobic signal
 CC sequence, a single 21-residue hydrophobic membrane-spanning domain,
 CC and a short, highly charged cytoplasmic domain. The extracellular
 CC portion contains six potential N-linked glycosylation sites and six
 CC Cys residues distributed among three C2 set Ig-related domains.
 CC A fusion protein of FCRI and a receptor ligand will be helpful to
 CC increase the potency of antibodies in therapy.

XX Sequence 344 AA;

Query Match 100.0%; Score 61; DB 13; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0097;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
 |||||
 Db 205 gnlvltlscetkl 216

RESULT 4
 ID R91439 standard; Protein; 344 AA.
 XX R91439;
 AC 30-OCT-1996 (first entry)
 DT XX
 DE Human FCRI (cDNA clone p98 product).

KM Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FCRI; Fc receptor.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 58 /note- "amino acid 58 is Leu in clone p135
 FT translated product"

PN US5506126-A.

XX 09-APR-1996.

PF 25-FEB-1988; 88US-0160416.

XX 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 18-OCT-1993; 93US-0139273.

XX (GEHO) GEN HOSPITAL CORP.

XX Aruffo A. Seed B;

PI WPI; 1996-200279/20.

DR N-PSDB; T14718.

XX Cloning of cDNA encoding cell surface antigen - useful for isolation

PT of diagnostic and therapeutic proteins

XX Example 10; Column 55-56; 79pp; English.

XX The amino acid sequence (R91439) of human FCRI was detd. from a

CC cDNA clone, p98 (T14718), obt'd. from a cDNA library using an

CC immunoselection cloning method. FCRI is a high affinity receptor

CC for the Fc portion of IgG, normally located on cell surfaces of

CC macrophages. 2 Other cDNA clones, p135 (T14717) and p90 (T14719),

CC coded for a variants of the FCRI sequence; the C-terminal sequence

CC of the p98 product is truncated compared with those of the p135

CC and p90 products (see also R91438 and W00859). A fusion protein
 CC of FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.

XX Sequence 344 AA;

Query Match 100.0%; Score 61; DB 17; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0097;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
 |||||
 Db 205 gnlvltlscetkl 216

RESULT 5
 ID Y96183 standard; Protein; 344 AA.
 XX Y96183;
 AC 19-DEC-2000 (first entry)
 DT XX
 DE Human macrophage-specific FCRI.

KM Macrophage; FCRI; cell surface antigen; human; immunoselection;
 KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 15 /note- "encoded by GCG"
 FT Misc-difference 38 /note- "encoded by ACC"

FT Misc-difference 50 /note- "encoded by CCG"

FT Misc-difference 51 /note- "encoded by ACC"

FT Misc-difference 55 /note- "encoded by CAC"

FT Misc-difference 56 /note- "encoded by TCC"

FT Misc-difference 60 /note- "encoded by CCC"

FT Misc-difference 64 /note- "encoded by CAG"

FT Misc-difference 82 /note- "encoded by CAA"

FT Misc-difference 116 /note- "encoded by GAA"

FT Misc-difference 117 /note- "encoded by GCA"

FT Misc-difference 122 /note- "encoded by TFC"

FT Misc-difference 123 /note- "encoded by ACC"

FT Misc-difference 126 /note- "encoded by CCG"

FT Misc-difference 129 /note- "encoded by CAT"

FT Misc-difference 134 /note- "encoded by AAT"

FT Misc-difference 136 /note- "encoded by GTT"

FT Misc-difference 139 /note- "encoded by CCA"

FT Misc-difference 140 /note- "encoded by AAT"

FT Misc-difference 213 /note- "encoded by AAT"

FT Misc-difference 216 /note= "encoded by CAA"
 FT /note= "encoded by TTC"
 FT Misc-difference 220 /note= "encoded by Cgr"
 FT Misc-difference 268 /note= "encoded by AAT"
 FT Misc-difference 305 /note= "encoded by GTG"
 FT Misc-difference 306 /note= "encoded by AAC"
 FT Misc-difference 332 /note= "encoded by GGT"
 FT Misc-difference 333 /note= "encoded by GGC"
 FT Misc-difference 338 /note= "encoded by CCT"
 XX
 PN US6111093-A.
 PD
 XX 29-AUG-2000.
 XX
 PF 28-OCT-1998; 98US-0181612.
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 82US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 23-MAR-1990; 90US-0488809.
 PR 13-JUL-1990; 90US-0553759.
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PL Stamenkovic I., Seed B;
 DR WPI: 2000-586382/55.
 DR N-PSDB; A50631.
 PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 PT useful for immunodiagnosis and immunotherapy of immune-mediated
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 PT diseases -
 PS Example 10; Column 53-55; 75pp; English.
 XX
 CC The present sequence is that of a human macrophage specific FCRI,
 CC as deduced from cDNA clone p135 (see A50631), which was isolated
 CC from a cDNA library expressed in COS cells using a novel method of
 CC the invention designed to isolate CSA nucleic acids. The method is
 CC based upon transient expression of a CSA in eukaryotic cells and
 CC physical selection of cells expressing the antigen by adhesion to
 CC (panning on) an antibody-coated substrate such as a culture dish.
 CC CSA nucleic acids isolated by the method of the invention, and the
 CC proteins they encode, are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders in
 CC animals, including humans. These disorders include asthma,
 CC immune-complex disease, amyloidosis, parasitic diseases or multiple
 CC sclerosis. FCRI is a high affinity receptor for the Fc portion of
 CC IgG, normally located on the cell surfaces of macrophages. The
 CC ability to interfere with such bonding, or to cause it to occur on
 CC surfaces other than macrophages, is useful in therapy. A fusion
 CC protein of FCRI and a receptor ligand will be helpful to increase
 CC the potencies of antibodies in therapy.

Query Match	100.0%	Score 61;	DB 21;	Length 344;
Best Local Similarity	100.0%	Pred. No. 0.0097;		
Matches 12; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1 GNLVLTSCETKL 12				

Db 205 gnlvtlscetk1 216

RESULT 6

ID R20811

AC R20811: R20811 standard; Protein; 374 AA.

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FCRI receptor encoded by clone p135.

KW Rapid immunoselection cloning technique; cell surface antigen; Immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN W09201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GEHO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amlot M;

DR WPI, 1992-056864/07.

DR N-PSDB; Q21178.

XX New CD53 cell surface antigen and DNA encoding it - for immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

XX This amino acid sequence was predicted from the cDNA sequence of cDNA clone p135. It differs from the sequence predicted from clones p98/X2 and p90 (see Q21179 and Q21180, respectively) at position 25; p135 encodes a Ser residue and the other two clones predict a Thr residue. At position 58, p135 predicts Leu and p90 predicts Val. Sequences predicted from all 3 clones show the typical features of a type I integral membrane protein and include a short hydrophobic signal sequence, a single 21-residue hydrophobic membrane-spanning domain, and a short, highly charged cytoplasmic domain. The extracellular portion contains six potential N-linked glycosylation sites and six Cys residues distributed among three C2 set Ig-related domains. A fusion protein of FCRI and a receptor ligand will be helpful to increase the potency of antibodies in therapy.

Sequence 374 AA;

```

Query Match Similarity      100.0% ; Score 61 ; DB 13 ; Length 374 ;
Best Local Similarity      100.0% ; Pred. No. 0.011 ;
Matches      12 ; Conservative      0 ; Mismatches      0 ; Indels      0 ; Gaps      0.
QY      1 GNLVLTSCETKL 12
          |||||
Db      205 gnlvltscetkl 216

RESULT      7
R22550
ID      R22550 standard ; Protein: 374 AA.
XX
XX      R22550 ;
XX
XX
DT      21-MAY-1992 (first entry)
XX

```


DE Human macrophage-specific FCRI receptor encoded by clone p90.
 XX
 KW Rapid immunoselection cloning technique; cell surface antigen;
 KW Immunodiagnosis; high affinity receptor.
 XX
 OS Homo sapiens.
 XX
 PN W09201049-A.
 PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90MO-US04986.
 XX
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO-) GEN HOSPITAL CORP.
 PI Seed B, Aruffo A, Amlot M;
 XX
 DR WPI: 1992-0556864/07.
 DR N-PSDB; Q21180.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX
 PS Example 10; Page 94a; 160pp; English.
 XX
 CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p90. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC and clone p90 predicts a Thr residue. At position 58, p90 predicts
 CC Val while clones p135 and p98/X2 (see Q21179) both predict Leu.
 CC Sequences predicted from all 3 clones show the typical features of
 CC a type I integral membrane protein and include a short hydrophobic
 CC signal sequence, a single 21-residue hydrophobic membrane-spanning
 CC domain, and a short, highly charged cytoplasmic domain. The
 CC extracellular portion contains six potential N-linked glycosylation
 CC sites and six Cys residues distributed among three C2 set Ig-related
 CC domains. A fusion protein of FCRI and a receptor ligand will be
 CC helpful to increase the potency of antibodies in therapy.
 CC
 SQ Sequence 374 AA;

Query Match 100.0%; Score 61; DB 13; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
 DB 205 gnlvtlscetkl 216

RESULT 8
 R91438
 ID R91438 standard; Protein; 374 AA.
 XX
 AC R91438;
 XX
 DT 30-OCT-1996 (first entry)
 XX
 DE Human FCRI (cDNA clone p135 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FCRI; Fc receptor.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 25 /note- "amino acid 25 is Thr in clone p90 and p98
 FT Misc-difference 58 translated products"

FT /note- "amino acid 58 is Val in p90 clone
 FT translated product"
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GEHO-) GEN HOSPITAL CORP.
 PI Aruffo A, Seed B;
 XX
 DR WPI: 1996-200279/20.
 DR N-PSDB; T14717.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX
 PS Example 10; Column 55-56; 79pp; English.
 XX
 CC The amino acid sequence (R91438) of human FCRI was detd. from a
 CC cDNA clone, p135 (T14717), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Another isolated cDNA clone, p90 (T14719), coded
 CC for a variant (W00859) of the FCRI sequence, and a third clone, p98
 CC (T14718), coded for an FCRI (R91439) having a different C-terminal
 CC sequence. A fusion protein of FCRI and a receptor ligand will
 CC be helpful in increasing the potency of antibodies in therapy.
 CC
 SQ Sequence 374 AA;

Query Match 100.0%; Score 61; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
 DB 205 gnlvtlscetkl 216

RESULT 9
 W00859
 ID W00859 standard; Protein; 374 AA.
 XX
 AC W00859;
 XX
 DT 30-OCT-1996 (first entry)
 XX
 DE Human FCRI (cDNA clone p90 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FCRI; Fc receptor.
 XX
 OS Homo sapiens.
 XX
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.

```

PR 13-JUL-1990: 90US-0553759.
PR 18-OCT-1993: 93US-0139273.
PA (GENO ) GEN HOSPITAL CORP.
PI Aruffo A, Seed B;
XX WPI: 1996-200279/20.
DR N-PSDB; T14719.
XX
XX Cloning of cDNA encoding cell surface antigen - useful for isolation
PT of diagnostic and therapeutic proteins
XX
XX Example 10; Column 55-56; 79pp; English.
PS
XX The amino acid sequence (W00859) of human FCRI was detd. from a
CC cDNA clone, p90 (T14719), obtd. from a cDNA library using an
CC immunoselection cloning method. FCRI is a high affinity receptor
CC for the Fc portion of IgG, normally located on cell surfaces of
CC macrophages. Other cDNA clones (see also T14717-18) coded for
CC variants (R91438-39) of the FCRI sequence. A fusion protein of
CC FCRI and a receptor ligand will be helpful in increasing the
CC potency of antibodies in therapy.
CC
XX Sequence 374 AA:
SO

```

Query Match 100.0%; Score 61; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 GNLVLTSCETKL 12
DB 205 gnlvltscetkl 216

```

RESULT 10
 ID W80448 standard; Protein; 374 AA.
 AC W80448;
 XX
 DT 07-JUN-1999 (first entry)
 DE Human Fc receptor I.
 KW Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 XX cloning.
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FT Misc-difference 2 /note= "encoded by TGG"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by CTG"
 FT Misc-difference 60 /note= "encoded by CCC"
 FT Misc-difference 77 /note= "encoded by AAT"
 FT Misc-difference 85 /note= "encoded by TCC"
 FT Misc-difference 99 /note= "encoded by CAA"
 FT Misc-difference 103 /note= "encoded by CCC"
 FT Misc-difference 141 /note= "encoded by GGC"
 FT Misc-difference 159 /note= "encoded by GGC"

```

FT /note= "encoded by AAC"
FT Misc-difference 171 /note= "encoded by ATG"
FT Misc-difference 176 /note= "encoded by GTC"
FT Misc-difference 256 /note= "encoded by GGC"
XX
XX US5830731-A.
XX
XX 03-NOV-1998.
XX
XX 21-MAY-1997; 97US-0861205.
XX
XX 01-DEC-1992; 92US-0983647.
XX 25-FEB-1988; 88US-0160416.
XX 13-JUL-1989; 89US-0379076.
XX 13-JUL-1990; 90US-0553759.
XX 21-MAY-1997; 97US-0861205.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX
XX Aruffo A, Seed B;
XX
XX WPI: 1996-609251/51.
XX N-PSDB; V63456.
XX
XX New cloning vector and polylinker - based on existing sequences for
PT efficient cloning and expression of mammalian cDNA(s), especially
PT human lymphocyte antigenic sequences
XX
XX Example 10; Column 53-54; 75pp; English.
PS
XX This is the amino acid sequence of human Fc receptor I (FCRI), as
CC deduced from cDNA clone p135 (see V63456) isolated using a rapid
CC immunoselection cloning method from a cDNA library expressed in COS
CC cells. The cDNA library was constructed from polyA RNA of cells
CC from a single patient undergoing extracorporeal interleukin-2
CC induction therapy. Clones p90 (see X07372) and p98 (see X07373)
CC were also obtained. A novel method for cloning cDNAs from mammalian
CC expression libraries is based on transient expression of an antigen
CC in eukaryotic cells and selection of cells expressing the antigen by
CC adhesion to an antibody-coated substrate. The method is useful for
CC the isolation and cloning of any protein which can be expressed and
CC transported to the cell surface membrane of a eukaryotic cell. It
CC has been used to clone genes (see V63442-63) encoding cell surface
CC antigens from mammalian lymphocytes (see W80440-55). The isolated
CC genes can be expressed in a prokaryotic or eukaryotic host cells to
CC produce the encoded protein. The invention also provides high
CC efficiency expression vectors (see V63441 and V63444) which allow
CC the generation of very large mammalian expression libraries. The
CC purified genes and proteins are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders of
CC animals, including humans.
XX
XX Sequence 374 AA:
SO

```

Query Match 100.0%; Score 61; DB 19; Length 374;
 Best Local Similarity ~ 100.0%; Pred. No. 0.011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 GNLVLTSCETKL 12
DB 205 gnlvltscetkl 216

```

RESULT 11
 ID W97833 standard; Protein; 374 AA.
 AC W97833;
 XX W97833;

```

XX 07-JUN-1999 (first entry)
DT Human Fc receptor I.
DE Human Fc receptor I.
XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
XX cloning.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 2 /note= "encoded by TGC"
XX FT Misc-difference 23 /note= "encoded by CTC"
XX FT Misc-difference 44 /note= "encoded by GAC"
XX FT Misc-difference 45 /note= "encoded by CTC"
XX FT Misc-difference 60 /note= "encoded by CCC"
XX FT Misc-difference 77 /note= "encoded by AAT"
XX FT Misc-difference 85 /note= "encoded by TCC"
XX FT Misc-difference 99 /note= "encoded by CAA"
XX FT Misc-difference 103 /note= "encoded by CCC"
XX FT Misc-difference 141 /note= "encoded by GGC"
XX FT Misc-difference 159 /note= "encoded by AAC"
XX FT Misc-difference 171 /note= "encoded by ATG"
XX FT Misc-difference 176 /note= "encoded by GTC"
XX FT Misc-difference 256 /note= "encoded by GGC"
XX FT Misc-difference 256 /note= "encoded by GGC"
XX
XX US5830731-A.
XX
XX 03-NOV-1998.
XX
XX 21-MAY-1997; 97US-0861205.
XX
XX 01-DEC-1992; 92US-0983647.
XX 25-FEB-1988; 88US-0160416.
XX 13-JUL-1989; 89US-0379076.
XX 13-JUL-1990; 90US-0553759.
XX 21-MAY-1997; 97US-0861205.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX
XX Aruffo A, Seed B;
XX
XX WPI; 1998-609251/51.
XX N-PSDB; X07372.
XX
XX New cloning vector and polylinker - based on existing sequences for
XX efficient cloning and expression of mammalian cDNA(s), especially
XX human lymphocyte antigenic sequences
XX
XX Example 10; Column 53-54; 75pp; English.
XX
XX This is the amino acid sequence of human Fc receptor I (FCRI), as
XX deduced from cDNA clone p90 (see X07372) isolated using a rapid
XX immunoselection cloning method from a cDNA library expressed in COS
XX cells. The cDNA library was constructed from polyA RNA of cells
XX from a single patient undergoing extracorporeal interleukin-2
XX induction therapy. Clones p135 (see V63456) and p98 (see X07373)
XX were also obtained. A novel method for cloning cDNAs from mammalian
XX expression libraries is based on transient expression of an antigen

```

```

CC in eukaryotic cells and selection of cells expressing the antigen by
CC adhesion to an antibody-coated substrate. The method is useful for
CC the isolation and cloning of any protein which can be expressed and
CC transported to the cell surface membrane of a eukaryotic cell. It
CC has been used to clone genes (see V63442-63) encoding cell surface
CC antigens from mammalian lymphocytes (see W80440-55). The isolated
CC genes can be expressed in a prokaryotic or eukaryotic host cells to
CC produce the encoded protein. The invention also provides high
CC efficiency expression vectors (see V63441 and V63444) which allow
CC the generation of very large mammalian expression libraries. The
CC purified genes and proteins are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders of
CC animals, including humans.
XX
XX Sequence 374 AA:
XX
XX Query Match 100.0%; Score 61; DB 19; Length 374;
XX Best Local Similarity 100.0%; Pred. No. 0.011;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GNLVTLSCETKL 12
XX Db 205 gnlvtlscetkl 216
XX
XX RESULT 12
XX ID W97834 standard; Protein; 374 AA.
XX AC W97834;
XX
XX 07-JUN-1999 (first entry)
XX
XX Human Fc receptor I.
XX
XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
XX cloning.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 2 /note= "encoded by TGC"
XX FT Misc-difference 23 /note= "encoded by CTC"
XX FT Misc-difference 44 /note= "encoded by GAC"
XX FT Misc-difference 45 /note= "encoded by CTC"
XX FT Misc-difference 60 /note= "encoded by CCC"
XX FT Misc-difference 77 /note= "encoded by AAT"
XX FT Misc-difference 85 /note= "encoded by TCC"
XX FT Misc-difference 99 /note= "encoded by CAA"
XX FT Misc-difference 103 /note= "encoded by CCC"
XX FT Misc-difference 141 /note= "encoded by GGC"
XX FT Misc-difference 159 /note= "encoded by AAC"
XX FT Misc-difference 171 /note= "encoded by ATG"
XX FT Misc-difference 176 /note= "encoded by GTC"
XX FT Misc-difference 256 /note= "encoded by GGC"
XX
XX US5830731-A.
XX

```

XX 03-NOV-1998.
 PD
 XX
 PF 21-MAY-1997; 97US-0861205.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 21-MAY-1997; 97US-0861205.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI, 1998-609251/51.
 DR N-PSDB; X07373.
 XX
 PT New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNA(s), especially
 PT human lymphocyte antigenic sequences
 PS
 XX Example 10; Column 53-54; 75pp; English.
 XX
 CC This is the amino acid sequence of human Fc receptor I (FCRI), as
 CC deduced from cDNA clone p98/X2 (see X07373) isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from polyA RNA of cells
 CC from a single patient undergoing extracorporeal Interleukin-2
 CC induction therapy. Clones p135 (see V63456) and p90 (see X07372)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnosis and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 CC
 XX
 SQ Sequence 374 AA;
 OY 1 GNLVLTSCETKL 12
 DB 205 gnlvltscetkl 216
 Y96134
 ID Y96134 standard; Protein; 374 AA;
 AC Y96134;
 DT 19-DEC-2000 (first entry)
 XX
 DE Human macrophage-specific FCRI.
 XX
 KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
 KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

XX OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 15
 FT /note= "encoded by GCG"
 FT Misc-difference 38
 FT /note= "encoded by ACC"
 FT Misc-difference 50
 FT /note= "encoded by CCG"
 FT Misc-difference 51
 FT /note= "encoded by ACC"
 FT Misc-difference 55
 FT /note= "encoded by CAC"
 FT Misc-difference 56
 FT /note= "encoded by TCC"
 FT Misc-difference 60
 FT /note= "encoded by CCC"
 FT Misc-difference 64
 FT /note= "encoded by CAG"
 FT Misc-difference 82
 FT /note= "encoded by CAA"
 FT Misc-difference 116
 FT /note= "encoded by GAA"
 FT Misc-difference 117
 FT /note= "encoded by GCA"
 FT Misc-difference 122
 FT /note= "encoded by TTC"
 FT Misc-difference 123
 FT /note= "encoded by ACC"
 FT Misc-difference 126
 FT /note= "encoded by CCG"
 FT Misc-difference 129
 FT /note= "encoded by CAT"
 FT Misc-difference 134
 FT /note= "encoded by AAT"
 FT Misc-difference 136
 FT /note= "encoded by GAT"
 FT Misc-difference 139
 FT /note= "encoded by CCA"
 FT Misc-difference 140
 FT /note= "encoded by AAT"
 FT Misc-difference 213
 FT /note= "encoded by CAA"
 FT Misc-difference 216
 FT /note= "encoded by TTC"
 FT Misc-difference 220
 FT /note= "encoded by CGT"
 FT Misc-difference 268
 FT /note= "encoded by AAT"
 FT Misc-difference 305
 FT /note= "encoded by GTG"
 FT Misc-difference 306
 FT /note= "encoded by AAC"
 FT Misc-difference 332
 FT /note= "encoded by GGT"
 XX
 PN US611093-A.
 XX
 PD 29-AUG-2000.
 XX
 PF 28-OCT-1998; 98US-0181612.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 23-MAR-1990; 90US-0498809.
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Stamenkovic I, Seed B;
 XX

DR N-PSDB; C77892.

XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
XX

PS Claim 11; Page 1739-1740; 2352pp; English.

CC C77607 to C78448 encode the human cancer associated proteins given in
CC B43398 to B44239. The proteins can have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic;
CC antiallergic; antirheumatic; antiarthritic; antiinflammatory;
CC antihypertoid; antiallergic; antibacterial; antiviral; dermatological;
CC neuroprotective; cardiac; thrombolytic; coagulant; nootropic;
CC vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and
CC polypeptides can be used for preventing, treating or ameliorating medical
CC conditions and diagnosing pathological conditions. Polynucleotides,
CC polypeptides, antibodies, agonists and antagonists from the present
CC invention may be used to treat immune disorders by activating or
CC inhibiting the proliferation, differentiation or mobilisation of immune
CC cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC allergic reactions, graft versus host disease and organ rejection,
CC modulate haemostatic or thrombolytic activity, modulate inflammation,
CC cancers, cardiovascular disorders, neurological disease and bacterial or
CC viral infections. The peptides, nucleotides, antibodies, agonists and
CC antagonists may be also be used in drug screens. C78449 to C78457 and
CC B44240 represent sequences used in the exemplification of the present
CC invention.
XX

SQ Sequence 399 AA;

Query Match 100.0%; Score 61; DB 21; Length 399;

Best Local Similarity 100.0%; Pred. No. 0.011; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNEVTLSCETKL 12
|||
Db 230 gnlvtlsceckl 241Search completed: June 4, 2001, 12:13:27
Job time: 200 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:28 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-22

Perfect score: 61

Sequence: 1 GNLVTLSCETKL 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/7A COMB pep: *
2: /cgn2_6/prodata/2/1aa/5B COMB pep: *
3: /cgn2_6/prodata/2/1aa/6A COMB pep: *
4: /cgn2_6/prodata/2/1aa/5B COMB pep: *
5: /cgn2_6/prodata/2/1aa/PCTUS COMB pep: *
6: /cgn2_6/prodata/2/1aa/backfile1.pcp: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	78.7	91	2	US-08-332-562A-94
2	37	60.7	94	3	US-08-297-395-10
3	37	60.7	113	2	US-08-466-860-7
4	37	60.7	113	3	US-08-472-040A-7
5	37	60.7	113	4	US-08-276-776-7
6	37	60.7	119	1	US-08-442-542-14
7	37	60.7	119	3	US-08-765-469-14
8	37	60.7	391	5	PCT-US95-15696-2
9	36	59.0	504	2	US-08-896-005-3
10	35	57.4	479	2	US-08-037-816A-20
11	35	57.4	479	2	US-08-037-816A-24
12	35	57.4	479	2	US-08-530-146-20
13	35	57.4	479	2	US-08-530-146-20
14	35	57.4	507	2	US-08-037-816A-16
15	35	57.4	507	2	US-08-037-816A-28
16	35	57.4	507	2	US-08-530-146-16
17	35	57.4	507	2	US-08-530-146-28
18	35	57.4	1130	2	US-08-519-547A-6
19	34	55.7	63	6	5284931-9
20	34	55.7	273	3	US-08-403-853-12
21	34	55.7	431	4	US-08-985-950-14
22	34	55.7	431	4	US-08-985-950-20
23	34	55.7	503	3	US-08-700-651-6
24	34	55.7	615	4	US-08-985-950-16
25	34	55.7	615	4	US-08-985-950-18
26	34	55.7	1479	4	US-08-840-062-2
27	33	54.1	20	1	US-08-218-025A-191

28	33	54.1	20	1	US-08-306-116A-25	Sequence 25, Appl
29	33	54.1	206	3	US-08-820-970-9	Sequence 9, Appl
30	33	54.1	278	2	US-08-432-016-5	Sequence 5, Appl
31	33	54.1	278	2	US-08-684-594-5	Sequence 5, Appl
32	33	54.1	306	1	US-08-424-682A-1	Sequence 1, Appl
33	33	54.1	319	1	US-08-633-148-4	Sequence 4, Appl
34	33	54.1	319	1	US-08-597-495B-22	Sequence 22, Appl
35	33	54.1	328	1	US-08-265-087-2	Sequence 2, Appl
36	33	54.1	328	1	US-08-186-529-2	Sequence 2, Appl
37	33	54.1	328	1	US-08-621-493-2	Sequence 2, Appl
38	33	54.1	328	1	US-08-640-386A-2	Sequence 2, Appl
39	33	54.1	328	2	US-08-965-688-2	Sequence 2, Appl
40	33	54.1	328	2	US-08-751-767A-4	Sequence 4, Appl
41	33	54.1	328	4	US-09-260-173-2	Sequence 2, Appl
42	33	54.1	340	2	US-08-633-148-2	Sequence 2, Appl
43	33	54.1	341	1	US-08-248-628A-2	Sequence 2, Appl
44	33	54.1	500	2	US-08-896-005-1	Sequence 1, Appl
45	33	54.1	519	2	US-08-751-767A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-332-562A-94
Sequence 94, Application US/08332562A
Patent No. 5985599
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONDAROVA, Lisa
APPLICANT: HOLETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRNA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO.: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-94
Query Match 78.7%; Score 48; DB 2; Length 91;

Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVLTSCETKL 12
1:|||||1111
DB 17 GSVLTSCETNCL 28

RESULT 2

US-08-297-395-10
Sequence 10, Application US/08297395A
Patent No. 6039947

GENERAL INFORMATION:

APPLICANT: Howard L. Weiner

APPLICANT: David A. Haller

TITLE OF INVENTION: PEPTIDES DERIVED FROM IMMUNODOMINANT

TITLE OF INVENTION: EPITOPES OF MYELIN BASIC PROTEIN

FILE REFERENCE: 1010/057230S3

CURRENT APPLICATION NUMBER: US/08/297,395A

CURRENT FILING DATE: 1994-08-11

EARLIER APPLICATION NUMBER: 08/059,189

EARLIER FILING DATE: 1993-05-06

EARLIER APPLICATION NUMBER: 07/502,559

EARLIER FILING DATE: 1990-03-30

EARLIER APPLICATION NUMBER: PCT/US88/02139

EARLIER FILING DATE: 1988-06-24

EARLIER APPLICATION NUMBER: 07/065,734

EARLIER FILING DATE: 1987-06-24

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 3.0;

SEQ ID NO 10

LENGTH: 94

TYPE: PRT

ORGANISM: Homo sapiens

US-08-297-395-10

Query Match 60.7%; Score 37; DB 3; Length 94;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GNLVLTSCETKL 12
1:|||||1111
DB 16 GQVLTSCETNCL 27

RESULT 3

US-08-466-860-7
Sequence 7, Application US/08466860
Patent No. 5985552

GENERAL INFORMATION:

APPLICANT: HOWELL, MARK D.

APPLICANT: BROSTOFF, STEVEN W.

APPLICANT: CARLO, DENNIS J.

TITLE OF INVENTION: VACCINATION AND METHODS AGAINST DISEASES

TITLE OF INVENTION: RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL

TITLE OF INVENTION: POPULATIONS

NUMBER OF SEQUENCES: 75

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CAMPBELL AND FLORES

STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700

CITY: SAN DIEGO

STATE: CALIFORNIA

COUNTRY: UNITED STATES

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,860

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/813,867

FILING DATE: 24-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: CAMPBELL, CATHRYN

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-IM 9107

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-535-9001

TELEFAX: 619-535-8949

INFORMATION FOR SEQ ID NO: 7;

SEQUENCE CHARACTERISTICS:

LENGTH: 113 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-08-466-860-7

Query Match 60.7%; Score 37; DB 2; Length 113;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GNLVLTSCETKL 12
1:|||||1111
DB 35 GQVLTSCETNCL 46

RESULT 4

US-08-472-040A-7
Sequence 7, Application US/08472040A
Patent No. 6030387

GENERAL INFORMATION:

APPLICANT: HOWELL, MARK D.

APPLICANT: BROSTOFF, STEVEN W.

APPLICANT: CARLO, DENNIS J.

TITLE OF INVENTION: VACCINATION AND METHODS AGAINST DISEASES

TITLE OF INVENTION: RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL

TITLE OF INVENTION: POPULATIONS

NUMBER OF SEQUENCES: 77

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CAMPBELL & FLORES LLP

STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700

CITY: SAN DIEGO

STATE: CALIFORNIA

COUNTRY: UNITED STATES

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,040A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: CAMPBELL, CATHRYN

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-IM 1641

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-535-9001

TELEFAX: 619-535-8949

INFORMATION FOR SEQ ID NO: 7;

SEQUENCE CHARACTERISTICS:

LENGTH: 113 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-08-472-040A-7

Query Match 60.7%; Score 37; DB 3; Length 113;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GNLVLTSCETKL 12
DB 35 GQNVLTSCQNTL 46

RESULT 5
US-08-276-776-7
Sequence 7, Application US/08276776
Patent No. 6207645
GENERAL INFORMATION:
APPLICANT: HOWELL, MARK D.
APPLICANT: BROSTOFF, STEVEN W.
APPLICANT: CARLO, DENNIS J.
TITLE OF INVENTION: VACCINATION AND METHODS AGAINST DISEASES
TITLE OF INVENTION: RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL
NUMBER OF SEQUENCES: 75
TITLE OF SEQUENCES: POPULATIONS
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: UNITED STATES
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,776
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/813,867
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, CATHERYN
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-1M 9107
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001
TELEFAX: 619-535-8949
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-276-776-7

Query Match 60.7%; Score 37; DB 4; Length 113;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GNLVLTSCETKL 12
DB 35 GQNVLTSCQNTL 46

RESULT 6
US-08-442-542-14
Sequence 14, Application US/08442542
Patent No. 5686600
GENERAL INFORMATION:

APPLICANT: Carozzi, Nadine B.
APPLICANT: Kozziel, Michael G.
TITLE OF INVENTION: Antibodies which Bind to Insect Gut
TITLE OF INVENTION: Proteins and their use
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,542
FILING DATE: 16-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/267,641
FILING DATE: 28-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: CGC 1750
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8615
TELEFAX: 919-541-8689
INFORMATION FOR SEQ. ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-442-542-14

Query Match 60.7%; Score 37; DB 1; Length 119;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNLVLTSCET 10
DB 15 GNSIKLSCET 24

RESULT 7
US-08-765-469-14
Sequence 14, Application US/08765469
Patent No. 6069301
GENERAL INFORMATION:
APPLICANT: Carozzi, Nadine B.
APPLICANT: Kozziel, Michael G.
TITLE OF INVENTION: Antibodies which Bind to Insect Gut
TITLE OF INVENTION: Proteins and their use
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,469

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/267,641
APPLICATION NUMBER: 08/267,641
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: SPULL, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: CGC 1750
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8615
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-469-14

Query Match 60.7%; Score 37; DB 3; Length 119;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GNLVLTSCET 10
11:11111
DB 15 GNSLKLSCET 24

RESULT 8
PCT-US95-15696-2
Sequence 2, Application PC/TUS9515696
GENERAL INFORMATION:
APPLICANT: President and Fellows of Harvard College
TITLE OF INVENTION: SINGLE CHAIN T-CELL RECEPTOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kevin M. Farrell, P.C.
STREET: P.O. Box 999
CITY: York Harbor
STATE: Maine
COUNTRY: USA
ZIP: 03911
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15696
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/349,915
FILING DATE: 06-DEC-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Farrell, Kevin M;
REGISTRATION NUMBER: 35,505
REFERENCE/DOCKET NUMBER: HU-9404 WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 207-363-0558
TELEFAX: 207-363-0528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 391 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15696-2

Query Match 60.7%; Score 37; DB 5; Length 391;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GNLVLTSCETKL 12
11111111
DB 169 GGNVLTSCENL 180

RESULT 9
US-08-896-005-3
Sequence 3, Application US/08896005
Patent No. 5854023
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,005
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0337 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 969078
US-08-896-005-3

Query Match 59.0%; Score 36; DB 2; Length 504;
Best Local Similarity 70.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GNLVLTSCET 10
11111111
DB 403 GNLVLTSCET 412

RESULT 10
US-08-037-816A-20
Sequence 20, Application US/08037816A
Patent No. 5865624
GENERAL INFORMATION:
APPLICANT: Hasel, Karl W.
APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED
TITLE OF INVENTION: THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES
TITLE OF INVENTION: THERETO
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM: /
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/037,816A
FILING DATE: 26-MAR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41190/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOPUI
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-037-816A-20

Query Match 57.4%; Score 35; DB 2; Length 479;

Best Local Similarity 54.5%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
DB 380 GNTITLPCRIR 390

RESULT 11

US-08-037-816A-24

Sequence 24, Application US/08037816A

Patent No. 5865624

GENERAL INFORMATION:

APPLICANT: Hasel, Karl W.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED

TITLE OF INVENTION: THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES

TITLE OF INVENTION: THERETO

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/037,816A

FILING DATE: 26-MAR-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41190/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOPUI
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-037-816A-24

Query Match 57.4%; Score 35; DB 2; Length 479;

Best Local Similarity 54.5%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
DB 380 GNTITLPCRIR 390

RESULT 12

US-08-530-146-20

Sequence 20, Application US/08530146

Patent No. 5865163

GENERAL INFORMATION:

APPLICANT: Hasel, Karl W.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED

TITLE OF INVENTION: THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES

TITLE OF INVENTION: THERETO

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/530,146

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/037,816

FILING DATE: 26-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41190/JPW/AJM

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

TELEX: 422523 COOPUI

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 479 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-530-146-20

Query Match 57.4%; Score 35; DB 2; Length 479;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
DB 380 GNLTLPCRIK 390

RESULT 13

US-08-530-146-24
Sequence 24, Application US/08530146

Patent No. 5886163

GENERAL INFORMATION:

APPLICANT: Hasel, Karl W.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED

TITLE OF INVENTION: THEREO, AND THERAPEUTIC AND PROPHYLACTIC USES

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/530,146

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/037,816

FILING DATE: 26-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41190/JPM/AJM

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

TELEX: 422523 COOPUI

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 479 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-530-146-24

QY 1 GNLVTLSCETK 11
DB 380 GNLTLPCRIK 390

RESULT 14

US-08-037-816A-16
Sequence 16, Application US/08037816A

Patent No. 5869624

GENERAL INFORMATION:

APPLICANT: Hasel, Karl W.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED

TITLE OF INVENTION: THEREO, AND THERAPEUTIC AND PROPHYLACTIC USES

TITLE OF INVENTION: THEREO

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/037,816A

FILING DATE: 26-MAR-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41190/JPM/AJM

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

TELEX: 422523 COOPUI

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 507 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-037-816A-16

QY 1 GNLVTLSCETK 11
DB 408 GNLTLPCRIK 418

RESULT 15
US-08-037-816A-28
Sequence 28, Application US/08037816A

Patent No. 5869624

GENERAL INFORMATION:

APPLICANT: Hasel, Karl W.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED

TITLE OF INVENTION: THEREO, AND THERAPEUTIC AND PROPHYLACTIC USES

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/037,816A

FILING DATE: 26-MAR-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41190/JPM/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOPUI
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-037-816A-28

Query Match 57.4%; Score 35; DB 2; Length 507;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 GNLYTISCETK 11
11:111111
Db 408 GMTTLPCKIK 418

Search completed: June 4, 2001, 12:14:29
Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:38 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-22

Sequence: 1 GNLVTLSCETKL 12

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

PIR 67: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	344	2 A41357	Fc gamma (Igg) rec
2	61	100.0	374	1 A39878	Fc gamma (Igg) rec
3	48	78.7	336	2 I48471	Fc gamma (Igg) rec
4	48	78.7	404	2 A46480	Fc gamma (Igg) rec
5	39	63.9	259	2 A71890	hydroxyethylthiazol
6	39	63.9	273	1 E64625	hydroxyethylthiazol
7	38	62.3	192	2 S33981	vlf protein - huma
8	38	62.3	314	2 A41544	developmental prot
9	38	62.3	314	2 A36195	developmental prot
10	38	62.3	584	2 T08678	hypothetical prote
11	37	60.7	113	2 A32578	T-cell receptor be
12	37	60.7	137	2 S57869	T-cell receptor Mb
13	37	60.7	142	2 S38388	T-cell receptor be
14	37	60.7	382	1 B69055	probable cell divi
15	37	60.7	1064	2 S74861	hypothetical prote
16	37	60.7	1170	2 T31971	hypothetical prote
17	36	59.0	64	2 A30569	Ig heavy chain V r
18	36	59.0	211	2 F71111	hypothetical prote
19	36	59.0	237	2 A42013	alpha-1-B-glycopro
20	36	59.0	430	2 T37198	hypothetical prote
21	36	59.0	487	2 T64033	adenosylhomocyste
22	36	59.0	527	1 S01302	adenosylhomocyste
23	36	59.0	669	2 A83731	DNA ligase (polyde
24	36	59.0	1797	2 T21889	hypothetical prote
25	36	59.0	1805	2 T21888	hypothetical prote
26	36	59.0	1829	2 T14280	RW1 protein - mous
27	36	59.0	3124	1 A40020	collagen alpha 1(X
28	35	57.4	26	2 A44036	collagen alpha 1(X
29	35	57.4	74	2 D81796	exodeoxyribonuclea

30	35	57.4	117	2 T07645	PEARL1 protein h
31	35	57.4	329	2 S00847	Ig gamma-2c chain
32	35	57.4	361	2 T45778	protein phosphatas
33	35	57.4	415	2 S51977	FUN9 protein - yea
34	35	57.4	424	2 T13906	citrate (sl)-synth
35	35	57.4	558	2 T15448	hypothetical prote
36	35	57.4	629	2 T34370	hypothetical prote
37	35	57.4	655	2 B65116	hypothetical 73.6
38	35	57.4	847	2 T09448	envelope glycoprot
39	35	57.4	847	2 S13289	env protein - huma
40	35	57.4	1130	2 A48843	MHC class II trans
41	35	57.4	5175	2 T20992	hypothetical prote
42	35	57.4	5198	2 T43290	hemisciental precurs
43	34	55.7	63	2 S09494	Ig heavy chain pre
44	34	55.7	156	2 A36922	of1 5' of acvB - A
45	34	55.7	168	2 A46482	T-cell surface gly

ALIGNMENTS

RESULT 1
Fc gamma (Igg) receptor I (high affinity) form b - human
N/Alternate names: CD64
C/Species: Homo sapiens (man)
C/Date: 03-Apr-1992 #sequence-revision 03-Apr-1992 #text-change 21-Jan-2000
C/Accession: A41357; S03019
R/Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A/Title: Isolation and expression of functional high-affinity Fc receptor complementa
A/Reference number: A41357; MUID:89100284
A/Accession: A41357
A/Molecule type: mRNA
A/Residues: 1-344 <ALL2>
A/Cross-references: GB:X14355; GB:M21090; NID:931333; PIDN:CAA32536.1; PID:931334
R/Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A/Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (
A/Reference number: S03018; MUID:89098339
A/Accession: S03019
A/Molecule type: mRNA
A/Residues: 1-344 <ALL2>
A/Cross-references: EMBL:X14355; NID:931333; PIDN:CAA32536.1; PID:931334
A/Note: the authors translated the codon ACT for residue 25 as Ala
C/Superfamily: Fc gamma receptor I; immunoglobulin homology
C/Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
P:117-170/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 61; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNLVTLSCETKL 12
Db 205 GNLVTLSCETKL 216

RESULT 2
Fc gamma (Igg) receptor I-A (high affinity) precursor - human
N/Alternate names: CD64
C/Species: Homo sapiens (man)
C/Date: 30-Dec-1991 #sequence-revision 06-Sep-1996 #text-change 21-Jul-2000
C/Accession: A39878; T70304; BA1357; S03018; I57525
R/van de Winkel, J.G.; J. Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
J. Biol. Chem. 266, 13449-13453, 1991
A/Title: Gene organization of the human high affinity receptor for IgG, Fc gammaRI (CD
A/Reference number: A39878; MUID:91302383
A/Accession: A39878
A/Molecule type: DNA
A/Residues: 1-374 <VAN>

```

A:Cross-references: GB:63830; GB:MG3835; NID:g180277; PIDN:CAA35678.1; PID:g180279
R:Porces, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
J. Clin. Invest. 90, 2102-2109, 1992
A>Title: Novel Fc gamma receptor 1 family gene products in human mononuclear cells.
A:Reference number: 153577; MUID:9305534
A:Accession: 170304
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-374 <RES>
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A>Title: Isolation and expression of functional high-affinity Fc receptor complementary
A:Reference number: A41357; MUID:89100284
A:Accession: B41357
A:Molecule type: mRNA
A:Residues: 1-24, 'S', 26-337, 'T', 339-374 <AL1>
A:Cross-references: GB:X14356; GB:M21091; NID:g31331; PIDN:CAA32537.1; PID:g31332
R:Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A>Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR
A:Reference number: 503018; MUID:89098339
A:Accession: 503018
A:Molecule type: mRNA
A:Residues: 1-24, 'S', 26-337, 'T', 339-374 <AL2>
A:Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332
R:Peretz, C.; Wietzerlin, J.; Benech, P.D.
Mol. Cell. Biol. 13, 2182-2192, 1993
A>Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
ism.
A:Reference number: 157525; MUID:93204964
A:Accession: 157525
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8 <RE2>
A:Cross-references: GB:557204; NID:g298692; PIDN:AAD13887.1; PID:g4261587
A:Accession: 557204
A:Genetics:
A:Gene: GDB:FCGR1A; CD64
A:Cross-references: GDB:135911; OMIM:146760
A:Map position: 1q21-1q21
C:Superfamily: Fc gamma receptor I; Immunoglobulin homology
C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane P
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-292/Domain: extracellular #status predicted <EXT>
F:117-170/Domain: immunoglobulin homology <IMM2>
F:293-313/Domain: transmembrane #status predicted <TM>
F:59,78,152,159,163,195,240/Binding site: carbohydrate (asn) (covalent) #status-predict
F:59,78,152,159,163,195,240/Binding site: carbohydrate (asn) (covalent) #status-predict
Query Match 100.0%; Score 61; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. NO. 0.0022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNLVTLSCETKL 12
Db 205 GNLVTLSCETKL 216
|||||
RESULT 3
148471 Fc gamma (IgG) receptor high affinity - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: 148471
R:Pinto, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarth, P.M.; Wicker, L.S.; Galf
Science 260, 695-698, 1993
A>Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for Ig
A:Reference number: 148471; MUID:93242399
A:Accession: 148471
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-336 <RES>
A:Cross-references: EMBL:X70980; NID:g311748; PIDN:CAA50311.1; PID:g311749

```

```

C:Superfamily:Fc gamma receptor I; immunoglobulin homology
C:Keywords: immunoglobulin receptor
F:128-180/Domain: immunoglobulin homology <IM>

Query Match      78.7%; Score 48; DB 2; Length 336;
Best Local Similarity 75.0%; Pred. No. 0.42;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
   |::|::|::|::|
Db 219 GSLVTLNCEINL 230

RESULT 4
A:Accession: A46480
Fc gamma (IlgG) receptor high affinity - mouse
N:Alternate names: high affinity IgG receptor
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
A:Accession: A46480; A43511
R:Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.
A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI a
A:Reference number: A46480; MUID:9216399
A:Accession: A46480
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-404 <OSM>
A:Note: sequence extracted from NCBI backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212,
R:Seers, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.
J:Immunol. 144, 371-378, 1990
A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for
A:Reference number: A43511; MUID:9011035
A:Accession: A43511
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-404 <SEA>
A:Cross-References: GB:M31314, NID:g200752, PIDN:AAA40056.1; PID:g200753
C:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: immunoglobulin receptor; transmembrane protein
F:127-179/Domain: immunoglobulin homology <IM>

Query Match      78.7%; Score 48; DB 2; Length 404;
Best Local Similarity 75.0%; Pred. No. 0.49;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12
   |::|::|::|::|
Db 214 GSLVTLNCEINL 225

RESULT 5
A:Accession: A71890
hydroxyethylthiazole kinase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jan-2000
A:Accession: A71890
R:Alm, R.A.; Ling, L.S.L.; Molt, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.;
Ires, C.; Gibson, R.; Meberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p
A:Reference number: A71800; MUID:99120557
A:Accession: A71890
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <ARN>
A:Cross-References: GB:AE001508; GB:AE001439; NID:g4155338; PIDN:AAD06352.1; PID:g415
A:Experimental source: strain J99
A:Genetics:
A:Gene: thm

```


C:Superfamily: hydroxyethylthiazole kinase; hydroxyethylthiazole kinase homology
F:2-189/Domain: hydroxyethylthiazole kinase homology <HMK>

Query Match 63.9%; Score 39; DB 2; Length 259;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11
I:|:|:|:|:|:|
Db 120 GSLVIGISCESK 130

RESULT 6
E64625
hydroxyethylthiazole kinase (EC 2.7.1.50) - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: E64625

R:Tomb, J.F.; White, O.; Keriavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.; Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467

A:Accession: E64625
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-273 <TOM>

A:Cross-references: GB:AE000595; GB:AE000511; NID:g2313969; PIDN:AMD07891.1; PID:g231397
C:Function:

A:Description: catalyzes the phosphorylation by ATP of 4-methyl-5-(2-hydroxyethyl)-thiaz
A:Pathway: thiamin biosynthesis
C:Superfamily: hydroxyethylthiazole kinase; hydroxyethylthiazole kinase homology
C:Keywords: phosphotransferase; thiamin biosynthesis
F:16-203/Domain: hydroxyethylthiazole kinase homology <HMK>

Query Match 63.9%; Score 39; DB 1; Length 273;
Best Local Similarity 63.6%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11
I:|:|:|:|:|:|
Db 134 GSLVIGISCESK 144

RESULT 7
S33981

vif protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
C:Accession: S33981

R:Carlini, F.
submitted to the EMBL Data Library, November 1991

A:Reference number: S33979
A:Accession: S33981
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-192 <CAR>
A:Cross-references: EMBL:Z11530; NID:g60192; PIDN:CA77623.1; PID:g60195
C:Superfamily: AIDS vif protein

Query Match 62.3%; Score 38; DB 2; Length 192;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 9
I:|:|:|:|:|:|
Db 126 GNLVTLSCETK 134

RESULT 8
A41544
developmental protein sina - fruit fly (Drosophila virilis)

C:Species: Drosophila virilis
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Feb-1997
C:Accession: A41544

R:Neufeld, T.P.; Cartnew, R.W.; Rubin, G.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10203-10207, 1991

A:Title: Evolution of gene position: Chromosomal arrangement and sequence comparison
A:Reference number: A41544; MUID:92052239
A:Accession: A41544

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-314 <NEU>

A:Cross-references: GB:M77281
C:Genetics:

A:Gene: FLYBase:FlyBase:FBgn0013142
A:Cross-references: FLYBase:FBgn0013142
C:Keywords: DNA binding; zinc finger

Query Match 62.3%; Score 38; DB 2; Length 314;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 12
I:|:|:|:|:|:|
Db 90 GNLVTLSCETK 101

RESULT 9
A36195
developmental protein sina - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
C:Date: 14-Dec-1990 #sequence_revision 13-Jan-1993 #text_change 24-Sep-1998
C:Accession: A36195

R:Cartnew, R.W.; Rubin, G.M.
Cell 63, 561-577, 1990

A:Title: seven in absentia, a gene required for specification of R7 cell fate in the
A:Reference number: A36195; MUID:91029488
A:Accession: A36195

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-314 <CAR>

A:Cross-references: GB:M38384; NID:g158466; PID:g158467
C:Genetics:

A:Gene: FLYBase:sina
A:Cross-references: FLYBase:FBgn0003410
C:Keywords: DNA binding; nucleus; zinc finger

Query Match 62.3%; Score 38; DB 2; Length 314;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 12
I:|:|:|:|:|:|
Db 90 GNLVTLSCETK 101

RESULT 10
T08678
hypothetical protein DKFZp56411922.1 - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
C:Accession: T08678

R:Hambitt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16469
A:Accession: T08678
A:Molecule type: mRNA
A:Residues: 1-584 <NMA>

A:Cross-references: EMBL:AL049946
 A:Experimental source: fetal brain; clone DKFZp56411922
 C:Genetics:
 A:Note: DKFZp56411922.1

Query Match 62.3%; Score 38; DB 2; Length 584;
 Best Local Similarity 54.5%; Pred. No. 42;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GNLVTLSCETK 11
 I:::|||||
 Db 117 GDVTVACEAK 127

RESULT 11
 A23578
 T-cell receptor beta chain precursor V region (HBV02) - human
 C:Species: Homo sapiens (man)
 C>Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 30-May-1997
 C:Accession: A32578
 R:Kimura, N.; Toyonaga, B.; Yoshikai, Y.; Du, R.P.; Mak, T.W.
 Eur. J. Immunol. 17, 375-383, 1987
 A:Title: Sequences and repertoire of the human T cell receptor alpha- and beta-chain var
 A:Reference number: A91263; MUID:87190670
 A:Accession: A32578
 A:Molecule type: mRNA
 A:Residues: 1-113 <KIM>
 A:Cross-references: GB:M27388
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor

Query Match 60.7%; Score 37; DB 2; Length 113;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 12
 I:::|||||
 Db 35 GQNVTLSCQNL 46

RESULT 12
 S57869
 T cell receptor Mh1 beta chain - human
 C:Species: Homo sapiens (man)
 C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 30-May-1997
 C:Accession: S57869
 R:Giegich, G.; Pette, M.; Meink, E.; Epplein, J.T.; Wekerle, H.; Hinkkanen, A.
 Eur. J. Immunol. 22, 753-758, 1992
 A:Title: Diversity of T cell receptor alpha and beta chain genes expressed by human T ce
 A:Reference number: S57869; MUID:92192051
 A:Accession: S57869
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-137 <GIE>
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor

Query Match 60.7%; Score 37; DB 2; Length 137;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 12
 I:::|||||
 Db 35 GQNVTLSCQNL 46

RESULT 13
 S38388
 T-cell receptor beta chain precursor - human (fragment)
 C:Species: Homo sapiens (man)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
 C:Accession: S38388
 R:Jones, R.; Moe, T.
 submitted to the EMBL Data Library, August 1993
 A:Description: Few V gene segments dominate the T cell receptor beta chain repertoire
 A:Reference number: S38386
 A:Accession: S38388
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-142 <JDR>
 A:Cross-references: EMBL:X74851; NID:9407429; PIDN:CAA52843.1; PID:9407430
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor

Query Match 60.7%; Score 37; DB 2; Length 142;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 12
 I:::|||||
 Db 35 GQNVTLSCQNL 46

RESULT 14
 B69055
 probable cell division control protein 6 - Methanobacterium thermoautotrophicum (stra
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: B69055
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514
 A:Accession: B69055
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-382 <MTD>
 A:Cross-references: GB:AE000903; GB:AE000666; NID:92622514; PIDN:AAB85889.1; PID:9262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1412
 C:Superfamily: Archaeoglobus fulgidus probable cell division control protein 6

Query Match 60.7%; Score 37; DB 1; Length 382;
 Best Local Similarity 54.5%; Pred. No. 43;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 NLVTLSCETK 12
 I:::|||||
 Db 364 NITDQCHTSL 374

RESULT 15
 S74861
 hypothetical protein slr1243 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C:Accession: S74861
 R:Raneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
 A:Reference number: S74322; MUID:97061201
 A:Accession: S74861
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1064 <KAN>

A:Cross-references: EMBL:D90909; GB:AB001339; NID:g1652844; PIDN:BA17822.1; PID:g165290
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C;superfamily: Synechocystis hypothetical protein slr1243

Query Match 60.7%; Score 37; DB 2; Length 1064;
 Best Local Similarity 45.5%; Pred. No. 1.1e+02;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11
 ||: ::|||:
 Db 623 GNIVNMACETR 633

Search completed: June 4, 2001, 12:15:39
 Job time: 272 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:01 ; Search time 37.56 seconds
(without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-22
Perfect score: 61
Sequence: 1 GNLVTLSCERKL 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	374	1	FCG1_HUMAN
2	48	78.7	404	1	P26151 mus musculi
3	39	63.9	259	1	THIM_HELPJ
4	39	63.9	259	1	THIM_HELPJ
5	38	62.3	314	1	SINA_DROME
6	38	62.3	314	1	SINA_DROME
7	38	62.3	314	1	SINA_DROME
8	37	60.7	1805	1	RM1_HUMAN
9	36	59.0	194	1	IL1B_RAT
10	36	59.0	457	1	CD4_SAISC
11	36	59.0	487	1	VPL_HAEN
12	36	59.0	504	1	SAH2_DROME
13	36	59.0	1829	1	RA1_MOUSE
14	36	59.0	3124	1	CA1L_CHICK
15	35	57.4	329	1	GCC_RAT
16	35	57.4	364	1	WN16_MOUSE
17	35	57.4	365	1	WN16_MOUSE
18	35	57.4	415	1	YAE2_YEAST
19	35	57.4	423	1	CISY_ACIAN
20	35	57.4	655	1	YHCP_ECOLI
21	35	57.4	669	1	NAH6_HUMAN
22	35	57.4	1130	1	C27A_HUMAN
23	35	57.4	3063	1	CA1C_MOUSE
24	35	57.4	193	1	IL18_HUMAN
25	34	55.7	196	1	EFTS_THETH
26	34	55.7	210	1	CD8B_HUMAN
27	34	55.7	210	1	CD8B_HUMAN
28	34	55.7	210	1	CD8B_HUMAN
29	34	55.7	219	1	EFTS_GUTTH
30	34	55.7	271	1	EFTS_MYCTU
31	34	55.7	271	1	TRPC_METTM
32	34	55.7	276	1	EFTS_MYCLE
33	34	55.7	413	1	PDK_DROME

34	34	55.7	474	1	ALBG_HUMAN	P04217 homo sapien
35	34	55.7	810	1	V2A_BMYV	P27462 broad bean
36	34	55.7	837	1	NOM2_HUMAN	O15384 homo sapien
37	34	55.7	1136	1	NOS2_CHICK	O90703 gallus gall
38	34	55.7	1694	1	SN_MOUSE	O62230 mus musculi
39	33	54.1	249	1	TSG_DROME	P54336 drosophila
40	33	54.1	274	1	YDA1_SCHPO	O10343 schizosacch
41	33	54.1	278	1	EFTS_STRCO	O31213 streptomyce
42	33	54.1	278	1	EFTS_STRRA	O95589 streptomyce
43	33	54.1	304	1	K2S2_HUMAN	P43631 homo sapien
44	33	54.1	304	1	K2S5_HUMAN	O14953 homo sapien
45	33	54.1	319	1	A33_HUMAN	O99795 homo sapien

ALIGNMENTS

RESULT	ID	FCG1_HUMAN	STANDARD	PRT	374 AA.
1	AC	P12314; P12315;			
DT	DT	01-OCT-1989 (Rel. 12, Created)			
DT	DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	DE	HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I) (CD64).			
GN	GN	FCGRIA OR FCGRI OR FCGI OR IGFR1.			
OS	OS	Homo sapiens (human).			
OC	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	OX	NCBI_TaxID=9606;			
RN	RN	[1]			
RP	RP	SEQUENCE FROM N.A.			
RX	RX	MEDLINE=89098339; PubMed=2974947;			
RA	RA	Allen J.M., Seed B.;			
RT	RT	"Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcRI)."			
RL	RL	Nucleic Acids Res. 16:11824-11824(1988).			
RN	RN	[2]			
RP	RP	SEQUENCE FROM N.A.			
RX	RX	MEDLINE=89100284; PubMed=2911749;			
RA	RA	Allen J.M., Seed B.;			
RT	RT	"Isolation and expression of functional high-affinity Fc receptor complementary DNAs."			
RL	RL	Science 243:378-381(1989).			
CC	CC	-1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH AFFINITY RECEPTOR.			
CC	CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	CC	-1- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.			
CC	CC	-1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.			
CC	CC	-1- DATABASE: NAME=PROW; NOTE=CD guide Cdb4 entry; WWW="http://www.ncbi.nlm.nih.gov/prow/cdb4.htm"			
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	CC	EMBL: X14356; CAA32537.1; -			
DR	DR	EMBL: X14355; CAA32536.1; -			
DR	DR	PIR: S03018; S03018.			
DR	DR	PIR: S03019; S03019.			
DR	DR	PIR: A41357; A41357.			
DR	DR	PIR: B41357; B41357.			
DR	DR	HSSP: P12319; 1ALT.			
DR	DR	MIM: 146760; -			
DR	DR	InterPro: IPR003006; -			
DR	DR	Pfam: PF00047; 1g; 3.			

KW 19g-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 Immunoglobulin domain; Alternative splicing; Polymorphism.

FT CHAIN 1 21
 FT 22 374
 FT 22 292
 FT TRANSMEM 293 313
 FT DOMAIN 314 374
 FT 36 92
 FT 117 175
 FT 205 267
 FT 43 85
 FT DISULFID 124 168
 FT DISULFID 212 260
 FT CARBOHYD 59 59
 FT CARBOHYD 78 78
 FT CARBOHYD 152 152
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT CARBOHYD 333 374
 FT 26 26
 FT 59 59
 FT 59 59
 FT 25 25
 FT 374 AA: 42605 MW: 2C2AAB103ECF16E6 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNLVITSCETKL 12
 DB 205 GNLVITSCETKL 216

RESULT 2
 FCGL_MOUSE
 ID FCGL_MOUSE STANDARD; PRT; 404 AA.
 AC P26151;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I).
 GN FCGR1 OR FCGL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90111035; PubMed-2136886;
 RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RT "Molecular cloning and expression of the mouse high affinity fc receptor for IgG";
 RL J. Immunol. 144:371-378(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92166399; PubMed-1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RT "Structure and mapping of the gene encoding mouse high affinity fc gamma RI and chromosomal location of the human fc gamma RI gene";
 RL J. Immunol. 148:1570-1575(1992).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 CC EMBL: M31314; AAA40056.1;
 CC DR PIR: A43511; A43511.
 CC DR PIR: A46480; A46480.
 CC DR HSSP: P12319; 1A1P.
 CC DR MGD: MGI:95498; Fcgr1.
 CC DR Interpro: IPR003006;
 CC DR Pfam: PF00047; 1g; 3.
 CC KW 19g-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 CC Immunoglobulin domain.
 CC FT CHAIN 1 24
 CC FT 25 404
 CC FT 25 297
 CC FT TRANSMEM 298 320
 CC FT DOMAIN 321 404
 CC FT 46 66
 CC FT 67 124
 CC FT 154 216
 CC FT 28 28
 CC FT CARBOHYD 48 48
 CC FT CARBOHYD 69 69
 CC FT CARBOHYD 168 168
 CC FT CARBOHYD 249 249
 CC FT 44887 MW: 1C4F0033842767E7 CRC64;

Query Match 78.7%; Score 48; DB 1; Length 404;
 Best Local Similarity 75.0%; Pred. No. 0.094;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNLVITSCETKL 12
 DB 214 GNLVITSCETKL 225

RESULT 3
 THIM_HELPJ
 ID THIM_HELPJ STANDARD; PRT; 259 AA.
 AC Q9RKZ9;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HYDROXYETHYLTHIAZOLE KINASE (EC 2.7.1.50) (4-METHYL-5-BETA-HYDROXYETHYLTHIAZOLE KINASE) (THZ KINASE) (TH KINASE).
 GN THIM OR JHP0783.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 OX NCBI_Taxid=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99120557; PubMed-9923682;
 RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Dolg P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nielsen W., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang O., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori";
 RL Nature 397:176-180(1999).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-METHYL-5-(2-HYDROXYETHYL)-THIAZOLE =
 CC ADP + 4-METHYL-5-(2-PHOSPHOETHYL)-THIAZOLE.
 CC -1- COFACTOR: MAGNESIUM (BY SIMILARITY).

CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE THZ KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AEO01508; AAD06352.1;
 CC InterPro: IPR000417;
 CC Pfam: PF02110; HK: 1.
 CC
 CC DR PRINTS: PR01099; HYERTHZKNASE.
 CC Thiamine biosynthesis; transferase; kinase; ATP-binding; Magnesium.
 CC METAL 86
 CC METAL 118 MAGNESIUM (BY SIMILARITY).
 CC FT ACT SITE 188 BASE (BY SIMILARITY).
 CC SEQUENCE 259 AA; 27646 MW; 2838CA8483C42000 CRC64;
 SO
 OY 1 GNLVLTSCETK 11
 Db 120 GSLVIGISCEK 130
 RESULT 4
 THIM_HELPY STANDARD; PRT; 259 AA.
 ID THIM_HELPY
 AC 025516;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HYDROXYETHYLTHIAZOLE KINASE (EC 2.7.1.50) (4-METHYL-5-BETA-
 GN THIM OR HP0845.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group.
 CC Helicobacter.
 CC NCBI_TaxID=210;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=26695 / ATCC 700392;
 CC MEDLINE=97394467; PubMed=9252185;
 CC
 CC RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 CC Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 CC Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 CC Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 CC McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 CC Berg D.E., Cooney J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 CC Colton R.D., Weidman J.M., Fujii C., Bowman C., Watney L., Wallin E.,
 CC Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 CC Venter J.C.;
 CC "The complete genome sequence of the gastric pathogen Helicobacter
 CC pylori".
 CC Nature 388:539-547(1997).
 CC
 CC -1- CATALYTIC ACTIVITY: ATP + 4-METHYL-5-(2-HYDROXYETHYL)-THIAZOLE =
 CC ADP + 4-METHYL-5-(2-PHOSPHOETHYL)-THIAZOLE.
 CC -1- COFACTOR: MAGNESIUM (BY SIMILARITY).
 CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE THZ KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AEO00595; AAD07891.1; ALT_INIT.
 CC TIGR: HP0845;
 CC InterPro: IPR000417;
 CC Pfam: PF02110; HK: 1.
 CC DR PRINTS: PR01099; HYERTHZKNASE.
 CC Thiamine biosynthesis; transferase; kinase; ATP-binding; Magnesium.
 CC METAL 86
 CC METAL 118 MAGNESIUM (BY SIMILARITY).
 CC FT ACT SITE 188 BASE (BY SIMILARITY).
 CC SEQUENCE 259 AA; 27514 MW; ACC0DC9B993F730 CRC64;
 SO
 OY 1 GNLVLTSCETK 11
 Db 120 GSLVIGISCEK 130
 RESULT 5
 SINA_DROME STANDARD; PRT; 314 AA.
 ID SINA_DROME
 AC P21461; Q9VVB0;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DEVELOPMENTAL PROTEIN SEVEN IN ABSENTIA.
 GN SINA OR CG9949.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CC NCBI_TaxID=7227;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=91029488; PubMed=2146028;
 CC
 CC RA Cartwright R.W., Rubin G.M.;
 CC "Seven in absentia, a gene required for specification of R7 cell fate
 CC in the Drosophila eye.";
 CC Cell 63:561-577(1990).
 CC
 CC -1- CATALYTIC ACTIVITY: ATP + 4-METHYL-5-(2-HYDROXYETHYL)-THIAZOLE =
 CC ADP + 4-METHYL-5-(2-PHOSPHOETHYL)-THIAZOLE.
 CC -1- COFACTOR: MAGNESIUM (BY SIMILARITY).
 CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE THZ KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Messaman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: REQUIRED FOR SPECIFICATION OF R7 PHOTORECEPTOR CELL
CC FATE IN THE DROSOPHILA EYE. POSSIBLY ACTS BY REGULATING GENE
CC EXPRESSION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: IN MANY ONMATIDIAL PRECURSOR CELLS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL: M38384; AAA28901.1;
DR EMBL: AE003526; AAF49403.1;
DR PIR: A36195; A36195.
DR Flybase: FBgn0003410; sina.
KW Developmental protein; Vision; Nuclear protein; Zinc-finger.
FT DOMAIN 14 21 POLY-ALA.
FT FT 46 55 POLY-SER.
FT DOMAIN 26 55 SER/THR-RICH.
FT SIMILAR 71 108 TO DICTYOSTELLUM DG17 (AA 25-67).
FT ZN.FING 73 107 POTENTIAL.
SQ SEQUENCE 314 AA: 33707 MW: 8587D5E8DA2F958 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 314;
Best Local Similarity 58.3%; Pred. No. 5.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 GNLVTLSCETKL 12
1:11:11:11:11
90 GHLVCVCSKSL 101

RESULT 6
SINA.DROVI STANDARD; PRT; 314 AA.
AC P29304;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DEVELOPMENTAL PROTEIN SEVEN IN ABSENTIA.
GN SINA.
OS *Drosophila virilis* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92052239; PubMed=1946441;
RA Neufeld T.P., Carthew R.W., Rubin G.M.;
RT "Evolution of gene position: chromosomal arrangement and sequence
RT comparison of the *Drosophila melanogaster* and *Drosophila virilis* sina
RT and Rha genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10203-10207(1991).
CC -1- FUNCTION: REQUIRED FOR SPECIFICATION OF R7 PHOTORECEPTOR CELL
CC FATE IN THE DROSOPHILA EYE. POSSIBLY ACTS BY REGULATING GENE

CC EXPRESSION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: IN MANY ONMATIDIAL PRECURSOR CELLS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL: M77282; AAA28899.1;
DR PIR: A41544; A41544.
DR Flybase: FBgn0013142; Dftr\sina.
KW Developmental protein; Vision; Nuclear protein; Zinc-finger.
FT DOMAIN 46 55 POLY-SER.
FT FT 27 55 SER/THR-RICH.
FT SIMILAR 71 108 TO DICTYOSTELLUM DG17 (AA 25-67).
FT ZN.FING 73 107 POTENTIAL.
SQ SEQUENCE 314 AA: 33804 MW: 66EAF8E814250EB CRC64;

Query Match 62.3%; Score 38; DB 1; Length 314;
Best Local Similarity 58.3%; Pred. No. 5.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 GNLVTLSCETKL 12
1:11:11:11:11
90 GHLVCVCSKSL 101

RESULT 7
C27A.MOUSE STANDARD; PRT; 1078 AA.
ID C27A.MOUSE
AC P79621; Q31115;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MHC CLASS II TRANSACTIVATOR CIITA.
DE MHC27A OR CIITA OR C27A.
GN MHC27A OR CIITA OR C27A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MOD; TISSUE=Spleen;
RX MEDLINE=97152412; PubMed=8995190;
RA Sims T.N., Elliott J.F., Ramassar V., Denney D.W. Jr., Halloran P.F.;
RT "Mouse class II transactivator: cDNA sequence and amino acid
RT comparison with the human class II transactivator.";
RL Immunogenetics 45:220-222(1997).
RN [2]
RP SEQUENCE OF 878-1020 FROM N.A.
RX STRAIN=BALB/C; TISSUE=Thymus;
RX MEDLINE=9618886; PubMed=8620527;
RA Panelli M.C., Wang E., Shen S., Schluter S.F., Bernstein R.M.,
RA Hersh E.M., Stopeck A., Gangavalli R., Barber J., Jolly D.,
RA Akporiaye E.T.;
RT "Interferon gamma (IFNgamma) gene transfer of an EMT6 tumor that is
RT poorly responsive to IFNgamma stimulation: increase in tumor
RT immunogenicity is accompanied by induction of a mouse class II
RT transactivator and class II MHC.";
RL Cancer Immunol. Immunother. 42:99-107(1996).
CC -1- FUNCTION: ESSENTIAL FOR TRANSCRIPTIONAL ACTIVITY OF THE HLA CLASS
CC II PROMOTER. ACTIVATION IS VIA THE PROXIMAL PROMOTER. NO DNA
CC BINDING OF IN VITRO TRANSLATED CIITA WAS DETECTED. MAY ACT IN A
CC COACTIVATOR-LIKE FASHION THROUGH PROTEIN-PROTEIN INTERACTIONS BY
CC CONTACTING FACTORS BINDING TO THE PROXIMAL MHC CLASS II PROMOTER,
CC TO ELEMENTS OF THE TRANSCRIPTION MACHINERY, OR BOTH. ALTERNATIVELY
CC IT MAY ACTIVATE HLA CLASS II TRANSCRIPTION BY MODIFYING PROTEINS
CC THAT BIND TO THE MHC CLASS II PROMOTER (BY SIMILARITY).


```
QY      1 GNLVTLSC 9
          ||| |||:
Db      1103 GNLNTLSCD 1111
```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE T-CELL SURFACE GLYCOPROTEIN CD4 PRECURSOR (T-CELL SURFACE ANTIGEN

```

DE GN T4/LEU-3).
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RA Tatsumi M., Hashimoto O.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: ACCESSORY PROTEIN FOR MHC CLASS-II ANTIGEN/T-CELL
CC RECEPTOR INTERACTION. MAY REGULATE T-CELL ACTIVATION.
CC -I- SUBUNIT: ASSOCIATES WITH P56-LCK (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I- SIMILARITY: BELONGS TO THE IMMUGLOBULIN SUPERFAMILY. CONTAINS
CC ONE V-LIKE AND 3 C2-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.sdb.ch).
CC -----
CC EMBL: D86588; BAA13131.1;
CC DR HSSP: P01730; IMBR.
CC DR InterPro: IPR000973;
CC DR InterPro: IPR003006;
CC Pfam: PF00047; Ig: 2.
CC DR PRINTS: PR00692; CD4TCANTIGEN.
CC Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell; MHC;
CC Signal; Lipoprotein; Palmitate.
CC KW CHAIN 1 25
CC FT SIGNAL 1 25 BY SIMILARITY.
CC FT CHAIN 26 457 T-CELL SURFACE GLYCOPROTEIN CD4.
CC FT DOMAIN 26 395 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 396 417 POTENTIAL.
CC FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 26 125 IG-LIKE V-TYPE DOMAIN.
CC FT DOMAIN 126 202 IG-LIKE C2-TYPE DOMAIN.
CC FT DOMAIN 203 316 IG-LIKE C2-TYPE DOMAIN.
CC FT DOMAIN 317 373 IG-LIKE C2-TYPE DOMAIN.
CC FT CARBOHYD 254 254 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC FT CARBOHYD 295 295 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC FT CARBOHYD 324 324 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC FT DISULFID 41 109 BY SIMILARITY.
CC FT DISULFID 155 184 BY SIMILARITY.
CC FT DISULFID 327 369 BY SIMILARITY.
CC FT LIPID 418 418 PALMITATE (BY SIMILARITY).
CC FT LIPID 421 421 PALMITATE (BY SIMILARITY).
CC SEQUENCE 457 AA; 50871 MW; 57ED6344005A015 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 457;
Best Local Similarity 58.3%; Pred.No. 20;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 GNVTLSCEFKL 12
| | | | | |
DB 34 GEVVELPCTSL 45

RESULT 11
VPL_HAEIN STANDARD: PRT: 487 AA.
AC P44233;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MU-LIKE PROPHAGE FLUMU TAIL SHEATH PROTEIN.
GN H11511.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

```

```

OC Haemophilus.
OX NCBI_TaxId=727;
RN RN
RP SEQUENCE FROM N.A.
RX STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: MAJOR COMPONENT OF THE TAIL (BY SIMILARITY).
CC -1- SIMILARITY: STRONG, TO PHAGE MU PROTEIN L.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32827; AAC23158.1; -.
DR DR
DR TIGR; H11511; -.
KW KW
KW Hypothetical protein.
SQ SEQUENCE 487 AA; 53143 MW; EB81ECA37C1278A7 CRC64;
-----
Query Match 59.0%; Score 36; DB 1; Length 487;
Best Local Similarity 54.5%; Pred. NO. 22;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 GNLVTLSCETK 11
DB 163 GNTLTTCRFK 173
-----
RESULT 12
ID SAH2_DROME STANDARD; PRT; 504 AA.
AC P50245; Q27587; Q9VE04;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE PUTATIVE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-
DE HOMOCYSTEINE HYDROLASE) (ADHCYASE).
GN AHCG89E OR PH200 OR CG8956.
GC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydrioidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN RN
RP SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RX MEDLINE=89030618; PubMed=2903049;
RA DeIoranzo M., Ali N., Saari G., Henry C., Wilcox M., Blenz M.;
RT "Evidence that the Abdominal-B r element function is conferred by a
RT trans-regulatory homeoprotein."
RL EMO J. 7:3223-3231(1988).
RN RN
RP SEQUENCE FROM N.A.
RP STRAIN=CANTON-S;
RX MEDLINE=95396803; PubMed=7667301;
RA Martin C.H., Mayeda C.A., Davis C.A., Ericsson C.L., Knafels J.D.,
RA Mathog D.R., Celniker S.E., Lewis E.B., Palazzolo M.J.;

```

Complete sequence of the bithorax complex of Drosophila.;
 Proc. Natl. Acad. Sci. U.S.A. 92:8398-8402(1995).
 RA SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-20196006: PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gockyne J.D.,
 Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 Abil J.F., Aghayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 Balles R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Mlshina N.V., Moadar C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
 Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reimert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 Switzas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 The genome sequence of Drosophila melanogaster.;
 Science 287:2185-2195(2000).
 RT Science 287:2185-2195(2000).
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O -
 CC ADENOSINE + L-HOMOCYSTEINE.
 CC -1- COFACTOR: NAD (BY SIMILARITY).
 CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EXTENDED GERM BAND EMBRYOS AND IN
 CC SOMATIC MESODERM, YOLK CELLS AND MIDGUT DURING LATER EMBRYONIC
 CC STAGES.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X13168; CAA13156.1; -
 CC EMBL: U31961; AAA84400.1; -
 CC EMBL: AE003715; AAF53567.1; -
 CC FlyBase: FBgn0015011; Ahcy89E.
 CC InterPro: IPR000043; -
 CC Pfam: PF00670; AdoHCYase_1; 1.
 CC PROSITE: PS00738; AdoHCYase_2; 1.
 CC PROSITE: PS00739; AdoHCYase_2; 1.
 CC HYDROLASE: NAD: One-carbon metabolism.
 CC NP_BIND 275 306 NAD (POTENTIAL).
 CC FT 127 127 T -> A (IN REF. 1).
 CC FT CONFLICT 484 504 PFKANYRLVTLTSLIASS -> LKPTTGMPLFFPPO

FT SEQUENCE 504 AA; 55442 MW; A3A4FER3894655BB CRC64;
 SQ
 Query Match 59.0%; Score 36; DB 1; Length 504;
 Best Local Similarity 70.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GNLVTLSCET 10
 DB 403 GRLVNLSCST 412
 RESULT 13
 RWL_MOUSE
 ID RWL_MOUSE STANDARD; PRT; 1829 AA.
 AC 070472;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE RWL PROTEIN.
 GN RWL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Brain;
 RA Wilkinson R., Fitter S., Tschirke D., Simmons A.;
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE IMMUNE RESPONSE TO VIRAL
 CC INFECTION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF060565; AAC1532.1; -
 CC DR MGD; MGI:1927110; Rwl.
 CC DR SEQUENCE 1829 AA; 200508 MW; DOB3E209257AFAB9 CRC64;
 SQ
 Query Match 59.0%; Score 36; DB 1; Length 1829;
 Best Local Similarity 77.8%; Pred. No. 83;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GNLVTLSCET 9
 DB 1131 GNLVTLSCET 1139
 RESULT 14
 CAIC_CHICK
 ID CAIC_CHICK STANDARD; PRT; 3124 AA.
 AC P13944; 004509;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR (FIBROCHIMERIN).
 GN COL12A1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RN RP STRAIN-WHITE LEGHORN;

FT CONFLICT 2803 2803 L -> F (IN REF. 2).
 FT CONFLICT 2977 2977 V -> F (IN REF. 2).
 FT CONFLICT 3075 3076 OP -> AG (IN REF. 3).
 SQ SEQUENCE 3124 AA: 340578 MM: 094285AFE7F346CF CRC64:

Search completed: June 4, 2001, 12:24:02
 Job time: 564 sec

Query Match
 Best Local Similarity 77.8%; Score 36; DB 1; Length 3124;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 NLVTLSCET 10
 DB 2499 NLVTLSCET 2507

RESULT 15

GCC_RAT STANDARD; PRT; 329 AA.
 ID GCC_RAT
 AC P20762;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-2C CHAIN C REGION.
 OS Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN
 RP
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88166903; PubMed=3127222;
 RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
 region cDNA: extensive homology to mouse gamma 3.";
 RL Eur. J. Immunol. 18:317-319(1988).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X07189; CAA30169.1;
 DR PIR: S00847; S00847.
 DR InterPro: IPR000495;
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 3.
 DR PROSITE: PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 113 HINGE.
 FT DOMAIN 114 222 CH2.
 FT DOMAIN 223 329 CH3.
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 143 203
 FT DISULFID 249 307
 SQ SEQUENCE 329 AA: 36571 MM: 5FCD7B7933850773 CRC64;

Query Match
 Best Local Similarity 57.4%; Score 35; DB 1; Length 329;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11
 DB 20 GNLVTLSCETK 30

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:23:21 ; Search time 107.68 Seconds
(without alignments)
13.062 Million cell updates/sec

```
Title: US-09-284-107-22
Perfect score: 61
Sequence: 1 GNLVTLSCETKL 12
```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum Match 1008
Listing first 45 summaries

```

1:  sp.archaea:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp.human:*
5:  sp.invertebrate:*
6:  sp.mammal:*
7:  sp.mhc:*
8:  sp.organelle:*
9:  sp.phage:*
10: sp.plant:*
11: sp.potent:*
12: sp.unclassified:*
13: sp.veterebrate:*
14: sp.virus:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution..

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	61	100.0	374	4	Q92663	Q92663 homo sapien
2	61	100.0	375	4	Q92495	Q92495 homo sapien
3	56	91.8	349	6	Q9MZT0	Q9mzt0 bos taurus
4	46	75.4	680	5	Q9Y1V6	Q9Y1V6 halocynthia
5	42	68.9	348	9	Q38131	Q38131 bacterioph
6	40	65.6	64	14	Q97753	Q97753 human immun
7	40	65.6	65	5	Q906A9	Q906A9 trypanosoma
8	39	63.9	428	10	Q9LUV7	Q9LUV7 arabidopsis
9	38	62.3	85	14	Q9QAY4	Q9qay4 yaba monkey
10	38	62.3	192	14	Q78242	Q78242 human immun
11	38	62.3	435	5	Q9V539	Q9V539 drosophila
12	38	62.3	584	4	Q9Y3Y8	Q9Y3Y8 homo sapien
13	38	62.3	1054	7	Q78109	Q78109 mus musculu
14	38	62.3	1155	7	Q9TPE1	Q9TPE1 mus musculu
15	38	62.3	2828	4	Q9NR99	Q9nr99 homo sapien
16	37	60.7	204	14	Q9JDC0	Q9jdc0 human immun
17	37	60.7	204	14	Q9JDE7	Q9jde7 human immun
18	37	60.7	285	4	Q43608	Q43608 homo sapien
19	37	60.7	382	1	Q27463	Q27463 methanobact

20	37	60.7	739	628260	cantis famili
21	37	60.7	1060	09Gxi3	09Gxi3 rattus norv
22	37	60.7	1064	2 P73770	P73770 synechocyst
23	37	60.7	1170	5 O16587	O16587 caenorhabd
24	37	60.7	1229	5 O9VT96	O9VT96 drosophila
25	37	60.7	1805	4 O92545	O92545 homo saplen
26	36	59.0	14	O97754	O97754 human immun
27	36	59.0	148	6 O9MZK7	O9MZK7 macaca mula
28	36	59.0	179	6 O9MZK9	O9MZK9 macaca mula
29	36	59.0	180	6 O9MZK8	O9MZK8 macaca mula
30	36	59.0	180	14 O9ORP6	O9ORP6 human immun
31	36	59.0	208	8 O9TM32	O9TM32 cyanidum c
32	36	59.0	211	1 O58394	O58394 pyrococcus
33	36	59.0	211	6 O28359	O28359 didelphis m
34	36	59.0	430	5 O44725	O44725 caenorhabd
35	36	59.0	476	2 O9LI32	O9LI32 streptomyce
36	36	59.0	547	2 O9N2U7	O9N2U7 caenorhabd
37	36	59.0	669	2 O9KFC37	O9KFC37 bacillus ha
38	36	59.0	763	5 O9XZD0	O9XZD0 drosophila
39	36	59.0	1829	11 O70472	O70472 mus musculu
40	36	59.0	1875	5 O93691	O93691 caenorhabd
41	35	57.4	26	6 O9TRM6	O9TRM6 bos taurus
42	35	57.4	63	14 O97756	O97756 human immun
43	35	57.4	64	14 O97755	O97755 human immun
44	35	57.4	64	14 O97751	O97751 human immun
45	35	57.4	64	14 O97752	O97752 human immun

ALIGNMENTS

```

RESULT      1
ID           092663
PRELIMINARY;          PRT;       374 AA.

AC           092663
DT           01-FEB-1997 (TREMBLrel_02, Created)
DT           01-FEB-1997 (TREMBLrel_02, Last sequence update)
DT           01-OCT-2000 (TREMBLrel_15, Last annotation update)
DE           FC GAMMA RECEPTOR I.
GN           A1.
OS           Hs sapiens (Human).
OC           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX           NCBI_TaxID=9606;
RN           (1)
RP           SEQUENCE FROM N.A.
RC           TISSUE=BLOOD;
RX           MELLIN=93055454; PubMed=1430234;
RA           Porjes A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RT           Kimberly R.P.;
RT           "Novel fc gamma receptor I family gene products in human mononuclear
RL           cells."
RL           J. Clin. Invest. 90:2102-2109(1992).
DR           EMBL; L03418; AAA36049.1; -
DR           HSSP; P12319; IALT.
DR           INTERPRO: IPR003006; -
DR           Pfam: PF00047; Ig; 3.1.
DR           PRODOM: PD00234; -; 1.
SQ           SEQUENCE 374 AA; 42632 MW; D33D59398CEA699 CRC64;

Query Match              100.0%; Score 61; Length 374;
Best Local Similarity    100.0%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY           1 GNLYTLSCETKL 12
            |||iiiiiii||
Db           205 GNLYTLSCETKL 216

RESULT      2
ID           092495
PRELIMINARY;          PRT;       375 AA.

```

AC 092495; 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR TYPE 1 (FC GAMMA RIB-FC GAMMA RECEPTOR)
 GN CD64 OR FC-GAMMA-RIB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Benech P.D., Sasstry K.N., Iyer R.R., Eichbaum O.G., Ravesh D.,
 RA Ezekowitz A.B.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-224 FROM N.A.
 RX MEDLINE=93018827; PubMed=1402657;
 RA Benech P.D., Sasstry K., Iyer R.R., Eichbaum O.G., Ravesh D.P.,
 RA Ezekowitz R.A.;
 RT *Definition of interferon gamma-response elements in a novel human Fc
 RT gamma receptor gene (Fc gamma R1b) and characterization of the gene
 RT structure.
 RL J. Exp. Med. 176:1115-1123(1992).
 DR EMBL; M91555; AAA58414.1; JOINED
 DR EMBL; M91550; AAA58414.1; JOINED
 DR EMBL; M91552; AAA58414.1; JOINED
 DR EMBL; M91553; AAA58414.1; JOINED
 DR EMBL; M91554; AAA58414.1; JOINED
 DR EMBL; M91554; AAA58414.1; JOINED
 DR EMBL; S45709; AAD13842.1; JOINED
 DR EMBL; S45707; AAD13842.1; JOINED
 DR EMBL; S45708; AAD13842.1; JOINED
 DR EMBL; S45705; AAD13842.1; JOINED
 DR HSSP; P12319; IALT.
 DR INTERPRO: IPR003006;
 DR PFAM: PF00047; 1g; 3.
 DR PRODOM: PD002534; -; 1.
 SO SEQUENCE 375 AA; 42881 MW; A84D464C70DD0F91 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 375;
 Best Local Similarity 100.0%; Pred. No. 0.0003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNLVLTSCETKL 12
 DB 206 GNLVLTSCETKL 217

RESULT 3
 ID 09M2T0 PRELIMINARY; PRT; 349 AA.
 AC 09M2T0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR I.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yan Y., Li X., Wang A., Zhang G.;
 RT "Molecular cloning and identification of full-length cDNA encoding
 RT high affinity Fc receptor for bovine IgG (Fc gamma RI).";
 RL Vet. Immunol. Immunopathol. 75:151-159(2000).
 DR EMBL; AF162866; AAF80477.1;
 KW Receptor.

SO SEQUENCE 349 AA; 39608 MW; D0B77B2EF9408C02 CRC64;

Query Match 91.8%; Score 56; DB 6; Length 349;
 Best Local Similarity 91.7%; Pred. No. 0.0028;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVLTSCETKL 12
 DB 205 GNLVLTSCETKL 216

RESULT 4
 ID 09Y1V6 PRELIMINARY; PRT; 680 AA.
 AC 09Y1V6;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE HRPT-3.
 GN HRPT-3.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Halocynthia.
 OX NCBI_TaxID=7729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sasakura Y., Ogasawara M., Makabe K.W.;
 RT "Two localization pathways of maternal RNAs at the posterior-vegetal
 RT cytoplasm in early ascidian embryos."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB029335; BAA81908.1;
 DR INTERPRO: IPR000050; -;
 DR PROSITE: PS01179; PID: 1.
 SO SEQUENCE 680 AA; 75525 MW; D44D68FB615D4670 CRC64;

Query Match 75.4%; Score 46; DB 5; Length 680;
 Best Local Similarity 90.0%; Pred. No. 0.57;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 NLVLTSCETKL 11
 DB 181 NLVLTSCETKL 190

RESULT 5
 ID 038131 PRELIMINARY; PRT; 348 AA.
 AC 038131;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
 DE STRUCTURAL PROTEIN.
 OS Bacteriophage rlt.
 OC Viruses.
 OX NCBI_TaxID=43685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96332668; PubMed=8730874;
 RA Nauta A., van Sinderen D., Karsens H., Smit E., Venema G., Kok J.;
 RT "Inducible gene expression mediated by a repressor-operator system
 RT isolated from Lactococcus lactis bacteriophage rlt.";
 RL Mol. Microbiol. 19:1331-1341(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96332669; PubMed=8730875;
 RA van Sinderen D., Karsens H., Kok J., Terpstra P., Rutgers M.H.,
 RA Venema G., Nauta A.;
 RT "Sequence analysis and molecular characterization of the temperate
 RT lactococcal bacteriophage rlt.";
 RL Mol. Microbiol. 19:1343-1355(1996).
 DR EMBL; U38906; AAB18720.1;

SO SEQUENCE 348 AA; 38233 MW; F6A9213E783F52C1 CRC64;

Query Match 68.9%; Score 42; DB 9; Length 348;

Best Local Similarity 72.7%; Pred. No. 1.9;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11

DB 267 GNLVTLSCETK 277

RESULT 6

ID 097753 PRELIMINARY; PRT; 64 AA.

AC 097753;

DT 01-FEB-1997 (TEMBLrel. 02, Created)

DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)

DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

GN ENV.

OS Human Immunodeficiency virus type 1.

OC Viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RP SEQUENCE FROM N.A.

RX MEDLINE=92092169; PubMed=1684385;

RA Pang S., Vinters H.V., Akashl T., O'Brien W.A., Chen I.S.,

"HIV-1 env sequence variation in brain tissue of patients with AIDS-

related neurologic disease."

RT J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).

RL EMBL: M75003; AAB08915.1;

DR INTERPRO: IPR000777;

DR PFAM: PF00516; GP120; 1.

KW Envelope protein.

FT NON_TER 1 64

FT NON_TER 64 64

SO SEQUENCE 64 AA; 7074 MW; A01822E1C3029C74 CRC64;

Query Match 65.6%; Score 40; DB 14; Length 64;

Best Local Similarity 63.6%; Pred. No. 0.88;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GNLVTLSCETK 11

DB 23 GNLVTLSCETK 33

RESULT 7

ID 0906A9 PRELIMINARY; PRT; 655 AA.

AC 0906A9;

DT 01-MAY-2000 (TEMBLrel. 13, Created)

DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)

DE NUCLEOLAR G-PROTEIN NOG1.

GN NOG1.

OS Trypanosoma brucei.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

OX NCBI_TaxID=5691;

RP SEQUENCE FROM N.A.

RC STRAIN=ISTAR1;

RA Park J.-H., Jensen B.C., Parsons M.;

"A novel nuclear G-protein conserved in eukaryotes."

RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF189284; AAF01061.1;

DR INTERPRO: IPR000765;

DR PRINTS: PRO0326; GTP10BG.

SO SEQUENCE 655 AA; 74797 MW; 3244326657BDA3E6 CRC64;

Query Match 65.6%; Score 40; DB 5; Length 655;

Best Local Similarity 66.7%; Pred. No. 8.9;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GNLVTLSCETKL 12

DB 331 GDLKTLACETKL 342

RESULT 8

ID 091UV7 PRELIMINARY; PRT; 428 AA.

AC 091UV7;

DT 01-OCT-2000 (TEMBLrel. 15, Created)

DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)

DE GB1AAC67359.1.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;

RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;

Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RN [12]

RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;

RA Nakamura Y.;

"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence

features of the regions of 4,504,864 bp covered by sixty P1 and TAC

clones."

RT DNA Res. 7:131-135(2000).

RL EMBL: AB022215; BAB01780.1;

SO SEQUENCE 428 AA; 50007 MW; A14268B7EB218FF6 CRC64;

Query Match 63.9%; Score 39; DB 10; Length 428;

Best Local Similarity 71.4%; Pred. No. 9.3;

Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 1 GNLVTLSC--ETKL 12

DB 265 GNLVTLSCVREKL 278

RESULT 9

ID 090AY4 PRELIMINARY; PRT; 85 AA.

AC 090AY4;

DT 01-MAY-2000 (TEMBLrel. 13, Created)

DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)

DE YB-ELIR.

GN YB-ELIR.

OS Yaba monkey tumor virus.

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Yabapoxvirus.

OX NCBI_TaxID=38804;

RP SEQUENCE FROM N.A.

RA Amano H., Kato K., Miyamura T.;

"Sequence analysis of Yaba virus DNA."

RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB025319; BA088869.1;

SO SEQUENCE 85 AA; 9386 MW; EF74954D554EBDEF CRC64;

Query Match 62.3%; Score 38; DB 14; Length 85;

Best Local Similarity 54.5%; Pred. No. 3;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 11
 DB 15 GNLVTLSCETK 25

RESULT 10
 ID 078242 PRELIMINARY: PRT: 192 AA.

AC 078242
 DT 01-NOV-1996 (TREMBLREL. 01, Created)
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLREL. 14, Last annotation update)
 DE VIF PROTEIN.
 GN VIF
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Carlini F., Federic M., Equestre M., Ricci S., Ratti G., Zibai Q.,
 RA Verani P., Rossi G.B.,
 RA J. Viral Diseases 1:40-55(1992).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89352106; PubMed=2765297;
 RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,
 RA Macchi B., Mangiano N., Verani P., Rossi G.,
 RT "Biologic and molecular characterization of producer and nonproducer
 RT clones from HT-78 cells infected with a patient HIV isolate.";
 RL AIDS Res. Hum. Retroviruses 5:385-396(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,
 RA Boiretti A., Saggio I., Verani P., Rossi G.,
 RA Int. J. Immunopharmacol. 3:17-23(1990).
 RL EMBL, J11530; CAA77623.1;
 DR INTERPRO: IPR000475;
 DR PFAM: PF00559; VIF; 1.
 DR PRINTS: PR00349; VIRIONINFCT.
 DR PRODOM: PD00063; -1.
 SQ SEQUENCE 192 AA; 22543 MW; 485E434FC92CCAD6 CRC64;

Query Match 62.3%; Score 38; DB 14; Length 192;
 Best Local Similarity 77.8%; Pred. No. 6.6;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNLVTLSC 9
 DB 126 GNLVTLSC 134

RESULT 11
 ID Q9V539 PRELIMINARY: PRT: 435 AA.

AC Q9V539;
 DT 01-MAY-2000 (TREMBLREL. 13, Created)
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
 DE BCDA:GH08789. PROTEIN.
 GN BCDA:GH08789.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burlis K.C., Bussan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Daveport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Stimpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL, AE003834; AAF58981.1;
 DR FLYBASE: FBgn0027561; BCDA:GH08789.
 DR INTERPRO: IPR001194;
 DR PFAM: PF02141; DENN; 1.
 SQ SEQUENCE 435 AA; 49783 MW; D9A1532EF5205E5B CRC64;

Query Match 62.3%; Score 38; DB 5; Length 435;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNLVTLSCETK 12
 DB 273 GEVTLNCDTKI 284

RESULT 12
 ID Q9Y3Y8 PRELIMINARY: PRT: 584 AA.

AC Q9Y3Y8;
 DT 01-NOV-1999 (TREMBLREL. 12, Created)
 DT 01-NOV-1999 (TREMBLREL. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
 DE HYPOHECTICAL 63.9 KDA PROTEIN (FRAGMENT).
 GN DKEZP56411922.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.,
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AL049946; CAB43220.1;
 DR HSSP: P56276; ITLK.
 DR INTERPRO: IPR003006;

DR PFAM: PF00047; 1q; 6.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 584 AA; 63918 MW; 9C75943922BADCA9 CRC64;

Query Match
 Best Local Similarity 62.3%; Score 38; DB 4; Length 584;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVLTSCETK 11
 1::1::1::1
 DB 117 GDVYVACEAK 127

RESULT 13
 ID 078109 PRELIMINARY; PRT; 1054 AA.
 AC 078109;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE MHC CLASS II TRANSACTIVATOR CIITA FORM IV.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX Muehlethaler-Mottet A., Otten L.A., Steimle V., Mach B.;
 RT "Expression of MHC class II molecules in different cellular and
 RT functional compartments is controlled by differential usage of
 RT multiple promoters of the transactivator CIITA."
 RL EMBO J. 16:2851-2860(1997).
 [2]
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX Otten L.A., Steimle V., Mach B.;
 RT "Equivalent expression of the Major Histocompatibility Complex class
 RT II genes through the various forms of the mouse transactivator
 RT CIITA."
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RA Otten L.A., Steimle V., Bontron S., Mach B.;
 RT "Equivalent expression of the Major Histocompatibility Complex class
 RT II genes through the various forms of the mouse transactivator
 RT CIITA."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF042159; AAC34367.1;
 DR EMBL; AF100710; AAF06839.1;
 KW MHC.
 SQ SEQUENCE 1054 AA; 116231 MW; EE446E7796C6FED8 CRC64;

Query Match
 Best Local Similarity 62.3%; Score 38; DB 7; Length 1054;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GNLVLTSCET 10
 1111111111
 DB 798 GNLVGLSCVT 807

RESULT 14
 ID 09TPP1 PRELIMINARY; PRT; 1155 AA.
 AC 09TPP1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE MHC CLASS II TRANSACTIVATOR CIITA FORM I.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=97327562; PubMed=9184229;
 RA Muehlethaler-Mottet A., Otten L.A., Steimle V., Mach B.;
 RT "Expression of MHC class II molecules in different cellular and
 RT functional compartments is controlled by differential usage of
 RT multiple promoters of the transactivator CIITA."
 RL EMBO J. 16:2851-2860(1997).
 [2]
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RA Otten L.A., Steimle V., Bontron S., Mach B.;
 RT "Equivalent expression of the Major Histocompatibility Complex class
 RT II genes through the various forms of the mouse transactivator
 RT CIITA."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF100709; AAF06838.1;
 KW MHC.
 SQ SEQUENCE 1155 AA; 127527 MW; F3FF05DEBFB0CE71 CRC64;

Query Match
 Best Local Similarity 62.3%; Score 38; DB 7; Length 1155;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GNLVLTSCET 10
 1111111111
 DB 899 GNLVGLSCVT 908

RESULT 15
 ID 09NR99 PRELIMINARY; PRT; 2828 AA.
 AC 09NR99;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ADLican.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Crowl R.M., Luk D.;
 RT "Identification of the gene encoding Adlican, a novel protein
 RT expressed in human arthritic tissues."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF245505; AAF86402.1;
 SQ SEQUENCE 2828 AA; 312292 MW; A18377D8554F1FE1 CRC64;

Query Match
 Best Local Similarity 54.5%; Score 38; DB 4; Length 2828;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNLVLTSCETK 11
 1::1::1::1
 DB 2361 GDVYVACEAK 2371

Search completed: June 4, 2001, 12:23:22
 Job time: 585 sec

Tue Jun 5 07:08:38 2001

us-09-284-107-22.rpt

Page 6

CC and p90 products (see also R91438 and W00859). A fusion protein
CC of FCRI and a receptor ligand will be helpful in increasing the
CC potency of antibodies in therapy.

CC Sequence 344 AA:

Query Match 100.0%; Score 61; DB 17; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIORPGQLYFS 12
Db 217 LIQRPGLYFS 228

RESULT 4
ID Y96183 standard; Protein: 344 AA.

AC Y96183;

DT 19-DEC-2000 (first entry)

DE Human macrophage-specific FCRI.

KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
KW immune disorder; infection; asthma; immune-complex disease;
KW amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 15 /note= "encoded by GCG"

FT Misc-difference 38 /note= "encoded by ACC"

FT Misc-difference 50 /note= "encoded by CCG"

FT Misc-difference 51 /note= "encoded by ACC"

FT Misc-difference 55 /note= "encoded by CAC"

FT Misc-difference 56 /note= "encoded by TCC"

FT Misc-difference 60 /note= "encoded by CCG"

FT Misc-difference 64 /note= "encoded by CAG"

FT Misc-difference 82 /note= "encoded by CAA"

FT Misc-difference 116 /note= "encoded by GAA"

FT Misc-difference 117 /note= "encoded by GCA"

FT Misc-difference 122 /note= "encoded by TTC"

FT Misc-difference 123 /note= "encoded by ACC"

FT Misc-difference 126 /note= "encoded by CCG"

FT Misc-difference 129 /note= "encoded by CAA"

FT Misc-difference 134 /note= "encoded by AAT"

FT Misc-difference 136 /note= "encoded by GGT"

FT Misc-difference 139 /note= "encoded by CCA"

FT Misc-difference 140 /note= "encoded by AAT"

FT Misc-difference 213

FT /note= "encoded by CAA"

FT Misc-difference 216 /note= "encoded by TTC"

FT Misc-difference 220 /note= "encoded by CCG"

FT Misc-difference 268 /note= "encoded by AAT"

FT Misc-difference 305 /note= "encoded by GTG"

FT Misc-difference 306 /note= "encoded by AAC"

FT Misc-difference 332 /note= "encoded by GGT"

FT Misc-difference 333 /note= "encoded by GCG"

FT Misc-difference 338 /note= "encoded by CCG"

PN US6111093-A.

PD 29-AUG-2000.

PF 28-OCT-1998; 98US-0181612.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 23-MAR-1990; 90US-0498809.

PR 13-JUL-1990; 90US-0553759.

XX (GENO) GEN HOSPITAL CORP.

XX Stamenkovic I, Seed B;

XX WPI: 2000-586382/55.

XX N-PSDB; A50631.

XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,

XX useful for immunodiagnosis and immunotherapy of immune-mediated

XX infections or disorders, e.g. asthma, immune-complex disease, parasitic

XX diseases

XX Example 10: Column 53-55; 75pp; English.

XX The present sequence is that of a human macrophage specific FCRI,

XX as deduced from cDNA clone p135 (see A50631), which was isolated

XX from a cDNA library expressed in COS cells using a novel method of

XX the invention designed to isolate CSA nucleic acids. The method is

XX based upon transient expression of a CSA in eukaryotic cells and

XX (panning on) an antibody-coated substrate such as a culture dish.

XX CSA nucleic acids isolated by the method of the invention, and the

XX proteins they encode, are useful for immunodiagnosis and

XX immunotherapeutic applications, including the diagnosis and

XX treatment of immune-mediated infections, diseases, and disorders in

XX animals, including humans. These disorders include asthma,

XX immune-complex disease, amyloidosis, parasitic diseases or multiple

XX sclerosis. FCRI is a high affinity receptor for the Fc portion of

XX IgG, normally located on the cell surfaces of macrophages. The

XX ability to interfere with such bonding, or to cause it to occur on

XX surfaces other than macrophages, is useful in therapy. A fusion

XX protein of FCRI and a receptor ligand will be helpful to increase

XX the potencies of antibodies in therapy.

CC Sequence 344 AA:

Query Match 100.0%; Score 61; DB 21; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIORPGQLYFS 12
IIIIIIIIIIIIIIII

Db 217 llqrpqlqlyfs 228

RESULT 5

ID R20811 standard; Protein: 374 AA.

AC R20811;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FCRI receptor encoded by clone p135.

DE Rapid immunoselection cloning technique; cell surface antigen;

DE Immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN WO9201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GEHO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amlot M;

DR WPI: 1992-056864/07.

DR N-PSDB; Q21180.

PT New CD53 cell surface antigen and DNA encoding it - for

PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

This amino acid sequence was predicted from the cDNA sequence of CC CDNA clone p135. It differs from the sequence predicted from CC clones p98/X2 and p90 (see Q21179 and Q21180, respectively) at CC position 25; p135 encodes a Ser residue and the other two clones CC predict a Thr residue. At position 58, p135 predicts Leu and p90 CC predicts Val. Sequences predicted from all 3 clones show the CC typical features of a type I integral membrane protein and include CC a short hydrophobic signal sequence, a single 21-residue CC hydrophobic membrane-spanning domain, and a short, highly charged CC cytoplasmic domain. The extracellular portion contains six CC potential N-linked glycosylation sites and six Cys residues CC distributed among three C2 set Ig-related domains. A fusion protein CC of FCRI and a receptor ligand will be helpful to increase the CC potency of antibodies in therapy.

SQ Sequence 374 AA:

Query Match 100.0%; Score 61; DB 13; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.003;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLQRPGLQLYFS 12

DB 217 llqrpqlqlyfs 228

RESULT 6

ID R22550 standard; Protein: 374 AA.

AC R22550;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FCRI receptor encoded by clone p90.

DE Rapid immunoselection cloning technique; cell surface antigen;

DE Immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN WO9201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GEHO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amlot M;

DR WPI: 1992-056864/07.

DR N-PSDB; Q21180.

PT New CD53 cell surface antigen and DNA encoding it - for

PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

This amino acid sequence was predicted from the cDNA sequence of CC CDNA clone p90. It differs from the sequence predicted from CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue CC and clone p90 predicts a Thr residue. At position 58, p90 predicts CC Val while clones p135 and p98/X2 (see Q21179) both predict Leu. CC Sequences predicted from all 3 clones show the typical features of CC a type I integral membrane protein and include a short hydrophobic CC signal sequence, a single 21-residue hydrophobic membrane-spanning CC domain, and a short, highly charged cytoplasmic domain. The CC extracellular portion contains six potential N-linked glycosylation CC sites and six Cys residues distributed among three C2 set Ig-related CC domains. A fusion protein of FCRI and a receptor ligand will be CC helpful to increase the potency of antibodies in therapy.

SQ Sequence 374 AA:

Query Match 100.0%; Score 61; DB 13; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.003;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLQRPGLQLYFS 12

DB 217 llqrpqlqlyfs 228

RESULT 7

ID R91438 standard; Protein: 374 AA.

AC R91438;

DT 30-OCT-1996 (first entry)

DE Human FCRI (cDNA clone p135 product).

DE Cell surface antigen; cloning; immunoselection; immunotherapy;

DE therapy; diagnosis; vector; FCRI; FC receptor.

OS Homo sapiens.

Key Location/Qualifiers

Misc-difference 25 /note= "amino acid 25 is Thr in clone p90 and p98

Misc-difference 58

/note- "amino acid 58 is Val in p90 clone translated product"

FT 13-JUL-1990; 90US-0553759.
 FT 18-OCT-1993; 93US-0139273.
 XX (GEHO) GEN HOSPITAL CORP.
 PA Aruffo A, Seed B;
 PI Aruffo A, Seed B;
 XX WPI, 1996-200279/20.
 XX N-PSDB; T14719.
 DR Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 CC
 XX Example 10; Column 55-56; 79pp; English.
 PS
 XX The amino acid sequence (W00859) of human FCRI was detd. from a
 CC cDNA clone, p15 (T14717), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Other cDNA clones (see also T14717-18) coded for
 CC variants (R91438-39) of the FCRI sequence. A fusion protein of
 CC FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.
 CC
 XX Sequence 374 AA;
 SO

Query Match 100.0%; Score 61; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDRPGLQLYFS 12
 DB 217 LLQRPGLQLYFS 228

RESULT 8

W00859
 ID W00859 standard; Protein; 374 AA.
 AC W00859;
 XX 30-OCT-1996 (first entry)
 DT Human FCRI (cDNA clone p90 product).
 DE Cell surface antigen; cloning; immunoselection; immunotherapy;
 XX therapy; diagnosis; vector; FCRI; Fc receptor.
 KM Homo sapiens.
 OS
 XX US5506126-A.
 PN 09-APR-1996.
 PD 25-FEB-1988; 88US-0160416.
 XX 01-DEC-1992; 92US-0983647.
 XX 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX (GEHO) GEN HOSPITAL CORP.
 PA Aruffo A, Seed B;
 PI Aruffo A, Seed B;
 XX WPI, 1996-200279/20.
 DR N-PSDB; T14719.
 DR Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 CC
 XX Example 10; Column 55-56; 79pp; English.
 PS
 XX The amino acid sequence (W00859) of human FCRI was detd. from a
 CC cDNA clone, p90 (T14719), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Other cDNA clones (see also T14717-18) coded for
 CC variants (R91438-39) of the FCRI sequence. A fusion protein of
 CC FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.
 CC
 XX Sequence 374 AA;
 SO

Query Match 100.0%; Score 61; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDRPGLQLYFS 12
 DB 217 LLQRPGLQLYFS 228

RESULT 9

W80448
 ID W80448 standard; Protein; 374 AA.
 AC W80448;
 XX 07-JUN-1999 (first entry)
 DT Human Fc receptor I.
 DE Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 XX cloning.
 KM Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH MISC-difference 2 /note- "encoded by TGG"
 FT MISC-difference 23 /note- "encoded by CTC"
 FT MISC-difference 44 /note- "encoded by GAC"
 FT MISC-difference 45 /note- "encoded by CTG"
 FT MISC-difference 60 /note- "encoded by CCC"
 FT MISC-difference 77 /note- "encoded by AAT"
 FT MISC-difference 85 /note- "encoded by TCC"
 FT MISC-difference 99 /note- "encoded by CAA"
 FT MISC-difference 103 /note- "encoded by CCC"
 FT MISC-difference 141 /note- "encoded by GGC"
 FT MISC-difference 159 /note- "encoded by GGC"

```

FT      /note= "encoded by AAC"
FT      Misc-difference 171
FT      /note= "encoded by ATG"
FT      Misc-difference 176
FT      /note= "encoded by GTC"
FT      Misc-difference 256
FT      /note= "encoded by GGC"
FT      /note= "encoded by GGC"
PN      US5830731-A.
XX      03-NOV-1998.
PD      21-MAY-1997; 97US-0861205.
XX      01-DEC-1992; 92US-0983647.
XX      25-FEB-1988; 88US-0160416.
XX      13-JUL-1989; 89US-0379076.
XX      13-JUL-1990; 90US-0553759.
XX      21-MAY-1997; 97US-0861205.
XX      (GEHO ) GEN HOSPITAL CORP.
PI      Aruffo A, Seed B;
XX      WPI; 1998-609251/51.
XX      N-PSDB; V63456.
XX      New cloning vector and polylinker - based on existing sequences for
XX      efficient cloning and expression of mammalian cDNA(s), especially
XX      human lymphocyte antigenic sequences
XX      Example 10; Column 53-54; 75pp; English.
XX      This is the amino acid sequence of human Fc receptor I (FcRI), as
XX      deduced from cDNA clone p135 (see V63456) isolated using a rapid
XX      immunoselection cloning method from a cDNA library expressed in COS
XX      cells. The cDNA library was constructed from polyA RNA of cells
XX      from a single patient undergoing extracorporeal interleukin-2
XX      induction therapy. Clones p90 (see X07372) and p98 (see X07373)
XX      were also obtained. A novel method for cloning cDNAs from mammalian
XX      expression libraries is based on transient expression of an antigen
XX      in eukaryotic cells and selection of cells expressing the antigen by
XX      adhesion to an antibody-coated substrate. The method is useful for
XX      the isolation and cloning of any protein which can be expressed and
XX      transported to the cell surface membrane of a eukaryotic cell. It
XX      has been used to clone genes (see V63442-63) encoding cell surface
XX      antigens from mammalian lymphocytes (see W80440-55). The isolated
XX      genes can be expressed in a prokaryotic or eukaryotic host cells to
XX      produce the encoded protein. The invention also provides high
XX      efficiency expression vectors (see V63441 and V63444) which allow
XX      the generation of very large mammalian expression libraries. The
XX      purified genes and proteins are useful for immunodiagnostic and
XX      immunotherapeutic applications, including the diagnosis and
XX      treatment of immune-mediated infections, diseases, and disorders of
XX      animals, including humans.
XX      Sequence 374 AA:
SQ

```

```

Query Match      100.0%; Score 61; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 LKORPGLQLYFS 12
        |||||
DB      217 LKRPGLQLYFS 228

```

```

RESULT 10
W97833
ID      W97833 standard; Protein; 374 AA.
XX
AC      W97833;

```

```

XX      07-JUN-1999 (first entry)
DF      Human Fc receptor I.
XX      Fc receptor I; FcRI; cell surface antigen; lymphocyte; human;
XX      cloning.
XX      Homo sapiens.
OS      Key Location/Qualifiers
XX      FT      Misc-difference 2 /note= "encoded by TGG"
XX      FT      Misc-difference 23 /note= "encoded by CTC"
XX      FT      Misc-difference 44 /note= "encoded by GAC"
XX      FT      Misc-difference 45 /note= "encoded by CTG"
XX      FT      Misc-difference 60 /note= "encoded by CCC"
XX      FT      Misc-difference 77 /note= "encoded by AAT"
XX      FT      Misc-difference 85 /note= "encoded by TCC"
XX      FT      Misc-difference 99 /note= "encoded by CAA"
XX      FT      Misc-difference 103 /note= "encoded by CCC"
XX      FT      Misc-difference 141 /note= "encoded by GGC"
XX      FT      Misc-difference 159 /note= "encoded by AAC"
XX      FT      Misc-difference 171 /note= "encoded by ATG"
XX      FT      Misc-difference 176 /note= "encoded by GTC"
XX      FT      Misc-difference 256 /note= "encoded by GGC"
XX      FT      /note= "encoded by GGC"
XX      US5830731-A.
XX      03-NOV-1998.
XX      21-MAY-1997; 97US-0861205.
XX      01-DEC-1992; 92US-0983647.
XX      25-FEB-1988; 88US-0160416.
XX      13-JUL-1989; 89US-0379076.
XX      13-JUL-1990; 90US-0553759.
XX      21-MAY-1997; 97US-0861205.
XX      (GEHO ) GEN HOSPITAL CORP.
XX      Aruffo A, Seed B;
XX      WPI; 1998-609251/51.
XX      N-PSDB; X07372.
XX      New cloning vector and polylinker - based on existing sequences for
XX      efficient cloning and expression of mammalian cDNA(s), especially
XX      human lymphocyte antigenic sequences
XX      Example 10; Column 53-54; 75pp; English.
XX      This is the amino acid sequence of human Fc receptor I (FcRI), as
XX      deduced from cDNA clone p90 (see X07372) isolated using a rapid
XX      immunoselection cloning method from a cDNA library expressed in COS
XX      cells. The cDNA library was constructed from polyA RNA of cells
XX      from a single patient undergoing extracorporeal interleukin-2
XX      induction therapy. Clones p135 (see V63456) and p98 (see X07373)
XX      were also obtained. A novel method for cloning cDNAs from mammalian
XX      expression libraries is based on transient expression of an antigen

```

CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.

SO Sequence 374 AA:

Query Match 100.0%; Score 61; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLORPGLQLYFS 12
 DB 217 Lldrpplqlyfs 228

RESULT 11
 ID W97834 standard; Protein: 374 AA.

XX W97834;

DT 07-JUN-1999 (first entry)

DE Human Fc receptor I.

XX Fc receptor I; FcRI; cell surface antigen; lymphocyte; human;
 KM cloning.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 2. /note= "encoded by TGG"

FT Misc-difference 23. /note= "encoded by CTC"

FT Misc-difference 44. /note= "encoded by GAC"

FT Misc-difference 45. /note= "encoded by CTC"

FT Misc-difference 60. /note= "encoded by CCC"

FT Misc-difference 77. /note= "encoded by AAT"

FT Misc-difference 85. /note= "encoded by TCC"

FT Misc-difference 99. /note= "encoded by CAA"

FT Misc-difference 103. /note= "encoded by CCC"

FT Misc-difference 141. /note= "encoded by GGC"

FT Misc-difference 159. /note= "encoded by AAC"

FT Misc-difference 171. /note= "encoded by ATG"

FT Misc-difference 176. /note= "encoded by GTC"

FT Misc-difference 256. /note= "encoded by GGG"

XX US5830731-A.

XX 03-NOV-1998.

XX 21-MAY-1997; 97US-0861205.

XX 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

XX 13-JUL-1989; 89US-0379076.

XX 13-JUL-1990; 90US-0553759.

XX 21-MAY-1997; 97US-0861205.

XX (GENO) GEN HOSPITAL CORP.

XX Aruffo A. Seed B;

XX WPI; 1998-609251/51.

XX N-PSDB; X07373.

XX New cloning vector and polylinker - based on existing sequences for

XX efficient cloning and expression of mammalian cDNA(s), especially

XX human lymphocyte antigenic sequences

XX Example 10; Column 53-54; 75pp; English.

XX This is the amino acid sequence of human Fc receptor I (FcRI), as

XX deduced from cDNA clone p98/X2 (see X07373) isolated using a rapid

XX immunoselection cloning method from a cDNA library expressed in COS

XX cells. The cDNA library was constructed from polyA RNA of cells

XX from a single patient undergoing extracorporeal interleukin-2

XX induction therapy. Clones p135 (see V63456) and p90 (see X07372)

XX were also obtained. A novel method for cloning cDNAs from mammalian

XX expression libraries is based on transient expression of an antigen

XX in eukaryotic cells and selection of cells expressing the antigen by

XX adhesion to an antibody-coated substrate. The method is useful for

XX the isolation and cloning of any protein which can be expressed and

XX transported to the cell surface membrane of a eukaryotic cell. It

XX has been used to clone genes (see V63442-63) encoding cell surface

XX antigens from mammalian lymphocytes (see W80440-55). The isolated

XX genes can be expressed in a prokaryotic or eukaryotic host cells to

XX produce the encoded protein. The invention also provides high

XX efficiency expression vectors (see V63441 and V63444) which allow

XX the generation of very large mammalian expression libraries. The

XX purified genes and proteins are useful for immunodiagnostic and

XX immunotherapeutic applications, including the diagnosis and

XX treatment of immune-mediated infections, diseases, and disorders of

XX animals, including humans.

XX Sequence 374 AA:

Query Match 100.0%; Score 61; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLORPGLQLYFS 12
 DB 217 Lldrpplqlyfs 228

RESULT 12
 ID Y96134 standard; Protein: 374 AA.

XX Y96134;

DT 19-DEC-2000 (first entry)

DE Human macrophage-specific FcRI.

XX Macrophage; FcRI; cell surface antigen; human; immunoselection;

XX panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;

XX immune disorder; infection; asthma; immune-complex disease;

XX amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 15 /note= "encoded by GCG"
XX FT Misc-difference 38 /note= "encoded by ACC"
XX FT Misc-difference 50 /note= "encoded by CCG"
XX FT Misc-difference 51 /note= "encoded by ACC"
XX FT Misc-difference 55 /note= "encoded by CAC"
XX FT Misc-difference 56 /note= "encoded by TCC"
XX FT Misc-difference 60 /note= "encoded by CCC"
XX FT Misc-difference 64 /note= "encoded by CAG"
XX FT Misc-difference 82 /note= "encoded by CAA"
XX FT Misc-difference 116 /note= "encoded by GAA"
XX FT Misc-difference 117 /note= "encoded by GCA"
XX FT Misc-difference 122 /note= "encoded by TTC"
XX FT Misc-difference 123 /note= "encoded by ACC"
XX FT Misc-difference 126 /note= "encoded by CCG"
XX FT Misc-difference 129 /note= "encoded by CAT"
XX FT Misc-difference 134 /note= "encoded by AAT"
XX FT Misc-difference 136 /note= "encoded by GTT"
XX FT Misc-difference 139 /note= "encoded by CCA"
XX FT Misc-difference 140 /note= "encoded by AAT"
XX FT Misc-difference 213 /note= "encoded by CAA"
XX FT Misc-difference 216 /note= "encoded by TTC"
XX FT Misc-difference 220 /note= "encoded by CGT"
XX FT Misc-difference 268 /note= "encoded by AAT"
XX FT Misc-difference 305 /note= "encoded by GTG"
XX FT Misc-difference 306 /note= "encoded by AAC"
XX FT Misc-difference 332 /note= "encoded by GGT"
XX FT Misc-difference 332 /note= "encoded by GGT"
XX PN US6111093-A.
XX PD 29-AUG-2000.
XX PF 28-OCT-1998; 98US-0181612.
XX PR 01-DEC-1992; 92US-0983647.
XX PR 25-FEB-1988; 88US-0160416.
XX PR 13-JUL-1989; 89US-0379076.
XX PR 23-MAR-1990; 90US-0498809.
XX PR 13-JUL-1990; 90US-0553759.
XX PA (GEHO ) GEN HOSPITAL CORP.
XX ST Stamenkovic I, Seed B;
XX PI
XX

```

```

DR WPI: 2000-586382/55.
DR N-PSDB; A50592.
XX
XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
XX useful for immunodiagnosis and immunotherapy of immune-mediated
XX infections or disorders, e.g. asthma, immune-complex disease, parasitic
XX diseases
XX
XX Example 10; Column 53-55; 75pp; English.
XX
XX The present sequence is that of a human macrophage specific FCRI,
XX as deduced from cDNA clone p135 (see A50592), which was isolated
XX from a cDNA library expressed in COS cells using a novel method of
XX the invention designed to isolate CSA nucleic acids. The method is
XX based upon transient expression of a CSA in eukaryotic cells and
XX (panning on) an antibody-coated substrate such as a culture dish.
XX CSA nucleic acids isolated by the method of the invention, and the
XX proteins they encode, are useful for immunodiagnosis and
XX immunotherapeutic applications, including the diagnosis and
XX treatment of immune-mediated infections, diseases, and disorders in
XX animals, including humans. These disorders include asthma,
XX immune-complex disease, amyloidosis, parasitic diseases or multiple
XX sclerosis. FCRI is a high affinity receptor for the FC portion of
XX IgG, normally located on the cell surfaces of macrophages. The
XX ability to interfere with such bonding, or to cause it to occur on
XX surfaces other than macrophages, is useful in therapy. A fusion
XX protein of FCRI and a receptor ligand will be helpful to increase
XX the potencies of antibodies in therapy.
XX
XX Sequence 374 AA:

```

```

Query Match 100.0%; Score 61; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LLORPGLQLYFS 12
DB 217 LIGRPGLQLYFS 228

```

```

RESULT 13
Y66226
ID Y96226 standard; Protein: 374 AA.
AC Y96226;
XX
XX 11-SEP-2000 (first entry)
DE Human high affinity Fc receptor, FcγγmARI.
XX
XX Human; high affinity Fc receptor; FcγγmARI; immunoglobulin;
XX infection; immune response; CD64; monocyte; macrophage; neutrophil;
XX eosinophil; HIV; IgG; immunosuppressive; antineoplastic; cytotoxic;
XX antiinflammatory; AIDS; SLE; multiple myeloma; rheumatoid arthritis;
XX systemic lupus erythematosus; tumour.
XX
XX Homo sapiens.
XX EP1006183-A1.
XX
XX 07-JUN-2000.
XX
XX 03-DEC-1998; 98EP-0122969.
XX
XX 03-DEC-1998; 98EP-0122969.
XX
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX WPI: 2000-367968/32.
XX DR N-PSDB; A27466.
XX

```

PT Novel Fc receptor lacking transmembrane domains, a signal peptide, and
PT glycosylation, useful for diagnosing and treating immune disorders and
PT cancer

PS Disclosure; Page 26-28; 60pp; English.

CC The present sequence is the human high affinity Fc receptor, FcgammaRI.
CC FcgammaRI is also known as CD64. Fc receptors play an important
CC role in defending the body against infections. First, pathogens are
CC opsonised by serum immunoglobulins. The resulting complex then binds to
CC cells expressing Fc receptors. FcgammaRI molecules are expressed by
CC monocytes and macrophages, but expression can also be induced on
CC neutrophils and eosinophils. Upon Fc receptor activation, immune effector
CC pathways are activated, leading to immune response. The present sequence
CC may be modified to produce recombinant versions. The recombinant Fc
CC receptor consist only of the extracellular portion of the receptor and
CC are not glycosylated i.e. they do not have transmembrane domains or
CC signal peptides. The recombinant proteins may be used in immunoassays to
CC determine the immune status of patients with chronic diseases of the
CC immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple
CC myeloma (MM), or Rheumatoid arthritis. In addition, pharmaceutical
CC compositions containing recombinant proteins may be used to treat or
CC prevent autoimmune diseases, allergies or tumours, especially AIDS,
CC rheumatoid arthritis or MM.

SO Sequence 374 AA;

Query Match 100.0%; Score 61; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGLQLYFS 12
DB 217 LIGRPGLQLYFS 228

RESULT 14

ID B43683 standard; Protein; 399 AA.

AC B43683;

DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1128.

KW Human: cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
KW antidiabetic; antistatic; antineumatic; antiarthritic; antiviral;
KW antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;
KW vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;
KW immune disorder; hematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW hemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.

KW Homo sapiens.

OS WO200055350-A1.

PN 21-SEP-2000.

PD 08-MAR-2000; 2000WO-US05882.

PF 12-MAR-1999; 99US-0124270.

PR (HUMA-) HUMAN GENE SCI INC.

PA Rosen CA, Ruben SM;

PI WPI; 2000-56753/55.

DR N-PSDB; C77892.

XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer

PS Claim 11; Page 1739-1740; 2352pp; English.

CC C77607 to C78448 encode the human cancer associated proteins given in
CC B43398 to B44239. The proteins can have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC cytostatic; proliferative; vulnerability; immunomodulator; antidiabetic;
CC antistatic; antineumatic; antiarthritic; antiinflammatory;
CC antihypertoid; antiallergic; antibacterial; antiviral; dermatologic;
CC neuroprotective; cardiant; thrombolytic; coagulant; nocotropic;
CC vasotropic; antiproliferative and angiogenic. The polynucleotides and
CC polypeptides can be used for preventing, treating or ameliorating medical
CC conditions and diagnosing pathological conditions. Polynucleotides,
CC polypeptides, antibodies, agonists and antagonists from the present
CC invention may be used to treat immune disorders by activating or
CC inhibiting the proliferation, differentiation or mobilisation of immune
CC cells, to treat disorders of hematopoietic cells, autoimmune disorders,
CC allergic reactions, graft versus host disease and organ rejection,
CC modulate haemostatic or thrombolytic activity, modulate inflammation,
CC cancers, cardiovascular disorders, neurological disease and bacterial or
CC viral infections. The peptides, nucleotides, antibodies, agonists and
CC antagonists may be also be used in drug screens. C78449 to C78457 and
CC B44240 represent sequences used in the exemplification of the present
CC invention.

SO Sequence 399 AA;

Query Match 100.0%; Score 61; DB 21; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLORPGLQLYFS 12
DB 242 LIGRPGLQLYFS 253

RESULT 15

ID R12428 standard; Protein; 410 AA.

AC R12428;

DT 14-AUG-1991 (first entry)

DE Hybrid Fc(gamma)RII/I receptor.

DE Immunoglobulin-binding molecule; Fc(gamma)RI; Fc(gamma)RII.

DE Key Location/Qualifiers

FT Region 1..203

FT Region /label=Fc(gamma)RII domains d1 and d2

FT Region /label=Fc(gamma)RI D3, transmembrane and

FT Region cytoplasmic domains

PN WO9106570-A.

PD 16-MAY-1991.

PF 25-OCT-1990; 90WO-AU00513.

PR 25-OCT-1989; 89AU-0007045.

PA (UYME-) UNIV MELBOURNE.

PI Hogarth PM, Hulet MD, Ierino FL, McKenzie IFC, Osman N;

PI WPI; 1991-164135/22.

DR N-PSDB; Q12079.

XX New immunoglobulin binding hybrid Fc receptor molecules - used to
PT control auto-immune diseases and allergic reactions and to regulate
PT antibody prodn.

XX Claim 10; fig 3B; 80pp; English.

XX This chimeric sequence is one example of a hybrid Fc receptor
CC of the invention. The hybrid Fc receptor retains the Ig binding
CC functions of the constituent receptor types. It can be used for
CC regulating antibody production in vivo and in plasmaphoresis to
CC remove immune complexes or pathological antibodies. Rabbit IgG
CC bound to cells transfected with the chimeric Fc(gamma) receptor.
CC See also Q12078.

XX Sequence 410 AA;

Query Match 91.8%; Score 56; DB 12; Length 410;

Best Local Similarity 91.7%; Pred. No. 0.026;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQRPGLQLYFS 12

Db 232 LQRPGLQLYFS 243

Search completed: June 4, 2001, 12:13:28
Job time: 201 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:29 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-23
Perfect score: 61
Sequence: 1 L1ORPGLQLXFS 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	91.8	91	2	US-08-332-562A-94
2	38	62.3	787	2	US-08-720-484A-4
3	38	62.3	787	4	US-08-953-823A-4
4	38	62.3	793	2	US-08-720-484A-2
5	38	62.3	793	4	US-08-953-823A-2
6	36	59.0	838	4	US-08-216-260-4
7	33	54.1	63	6	5304637-7
8	33	54.1	150	6	5304637-1
9	33	54.1	152	1	US-08-284-393B-12
10	33	54.1	152	5	PCRT-US95-08950-12
11	33	54.1	628	2	US-07-952-853-22
12	33	54.1	628	2	US-08-914-848-22
13	32	52.5	123	3	US-08-840-316-3
14	32	52.5	123	3	US-08-478-507-9
15	32	52.5	123	4	US-08-809-523-3
16	32	52.5	123	5	PCRT-US93-08849A-3
17	32	52.5	123	5	PCRT-US93-08849-3
18	32	52.5	124	5	US-08-240-049B-19
19	32	52.5	135	4	PCRT-US95-13703-21
20	32	52.5	135	4	US-09-293-395-2
21	32	52.5	135	4	US-09-293-395-5
22	32	52.5	135	4	US-09-293-395-13
23	32	52.5	142	3	US-09-053-197A-33
24	32	52.5	269	2	US-07-857-224B-48
25	32	52.5	271	2	US-07-857-224B-47
26	32	52.5	683	3	US-08-630-916A-46
27	32	52.5	854	2	US-09-070-060-4

28	32	52.5	854	3	US-09-357-746-4	Sequence 4, Appl
29	32	52.5	906	3	US-08-630-916A-48	Sequence 48, Appl
30	32	52.5	1053	3	US-08-613-009A-8	Sequence 8, Appl
31	32	52.5	1070	3	US-08-613-009A-11	Sequence 11, Appl
32	32	52.5	1074	3	US-08-613-009A-7	Sequence 7, Appl
33	32	52.5	1076	2	US-08-867-941-19	Sequence 19, Appl
34	32	52.5	1076	4	US-09-074-658-19	Sequence 19, Appl
35	32	52.5	1447	5	PCRT-US94-05277-2	Sequence 2, Appl
36	32	52.5	2004	1	US-08-375-709-15	Sequence 15, Appl
37	32	52.5	2004	1	US-08-752-929-15	Sequence 15, Appl
38	32	52.5	2004	4	US-09-090-793-9	Sequence 9, Appl
39	32	52.5	2500	3	US-08-801-263A-2	Sequence 2, Appl
40	32	52.5	2500	3	US-09-102-248-2	Sequence 2, Appl
41	32	52.5	2512	2	US-08-801-263A-9	Sequence 9, Appl
42	32	52.5	2512	3	US-09-102-248-9	Sequence 9, Appl
43	32	52.5	2517	2	US-08-801-263A-5	Sequence 5, Appl
44	32	52.5	2517	3	US-09-102-248-5	Sequence 5, Appl
45	32	52.5	2584	3	US-08-936-135-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-332-562A-94
Sequence 94, Application US/08332562A
Patent No. 5985599
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIRBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-94
Query Match 91.8%, Score 56, DB 2, Length 91:

Best Local Similarity 91.7%; Pred. No. 0.0024;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 L10RPGQLYFS 12
DB 29 L10RPGQLYFS 40

RESULT 2

US-08-720-484A-4
Sequence 4, Application US/08720484A
Patent No. 5990281
GENERAL INFORMATION:
APPLICANT: Desauvage, Frederic
APPLICANT: Rosenthal, Arnon
APPLICANT: Stone, Donna
TITLE OF INVENTION: Vertebrate Smoothened Proteins
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,484A
FILING DATE: 30-Sep-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1050
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-720-484A-4

Query Match 62.3%; Score 38; DB 2; Length 787;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 L10RPGQLYFS 12
DB 579 L10RPGQLYFS 590

RESULT 3

US-08-953-823A-4
Sequence 4, Application US/08953823A
Patent No. 6136958
GENERAL INFORMATION:
APPLICANT: Desauvage, Frederic
APPLICANT: Rosenthal, Arnon
APPLICANT: Stone, Donna
TITLE OF INVENTION: Vertebrate Smoothened Proteins
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California

COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/953,823A
FILING DATE: 30-Sep-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027070
FILING DATE: 30-Sep-1996
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1050R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-953-823A-4

Query Match 62.3%; Score 38; DB 4; Length 787;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 L10RPGQLYFS 12
DB 579 L10RPGQLYFS 590

RESULT 4

US-08-720-484A-2
Sequence 2, Application US/08720484A
Patent No. 5990281
GENERAL INFORMATION:
APPLICANT: Desauvage, Frederic
APPLICANT: Rosenthal, Arnon
APPLICANT: Stone, Donna
TITLE OF INVENTION: Vertebrate Smoothened Proteins
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,484A
FILING DATE: 30-Sep-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1050
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 793 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-720-484A-2

Query Match 62.3%; Score 38; DB 2; Length 793;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LQRPGLQLYFS 12
Db 583 LQNPQGLSFS 594

RESULT 5
US-08-953-823A-2
Sequence 2, Application US/08953823A
Patent No. 6136958
GENERAL INFORMATION:
APPLICANT: Desauvage, Frederic
APPLICANT: Rosenthal, Arnon
TITLE OF INVENTION: Vertebrate Smoothed Proteins
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/953,823A
FILING DATE: 30-Sep-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027070
FILING DATE: 30-Sep-1996
ATTORNEY/AGENT INFORMATION:
NAME: Syvoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1050R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-953-823A-2

Query Match 62.3%; Score 38; DB 4; Length 793;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LQRPGLQLYFS 12
Db 583 LQNPQGLSFS 594

RESULT 6
US-08-216-260-4
Sequence 4, Application US/08216260
Patent No. 5837261
GENERAL INFORMATION:

APPLICANT: Inglis, Stephen C.
APPLICANT: Bournelli, Michael E.G.
APPLICANT: Minson, Anthony C.
TITLE OF INVENTION: VIRAL VACCINES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,260
FILING DATE: 21-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9305710.7
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9324964.7
FILING DATE: 06-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/030,073
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/168,643
FILING DATE: 16-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-59886/RMD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 838 amino acids
TYPE: amino acid
STRANDEDNESS: both
TOPOLOGY: linear
US-08-216-260-4

Query Match 59.0%; Score 36; DB 2; Length 838;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LQRPGLQLYFS 12
Db 227 LRSRGRVYFS 238

RESULT 7
Patent No. 5304637
APPLICANT: DORSSERS, LAMBERTUS C.J.; WAGEMAKER, GERARD; VOS,
VONNE J.; VAN LEEN, ROBERT W.; PERSOON, MARIA L.N.
TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF HUMAN
INTERLEUKIN-3 AND MOTILINS THEREOF
NUMBER OF SEQUENCES: 30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/494,182
FILING DATE: 13-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 249,184

FILING DATE: 16-AUG-1988
SEQ ID NO: 7
LENGTH: 63
5304637-7

Query Match 54.1%; Score 33; DB 6; Length 63;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLORPGLQ 8
11111111
Db 35 LTVRPGLO 42

RESULT 8
5304637-1
PATENT NO. 5304637
APPLICANT: DORSSERS, LAMBERTUS C.J.; WAGEMAKER, GERARD; VOS,
VONNE J.; VAN LEEN, ROBERT W.; PERSOON, MARIA L.N.
TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF HUMAN
INTERLEUKIN-3 AND MUTINS THEREOF
NUMBER OF SEQUENCES: 30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/494,182
FILING DATE: 13-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 249,184
FILING DATE: 16-AUG-1988
SEQ ID NO: 1
LENGTH: 150
5304637-1

Query Match 54.1%; Score 33; DB 6; Length 150;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLORPGLQ 8
11111111
Db 12 LTVRPGLO 19

RESULT 9
US-08-284-393B-12
Sequence 12, Application US/08284393B
Patent No. 5696234
GENERAL INFORMATION:
APPLICANT: Zurawski, Sandra M.
APPLICANT: Zurawski, Gerard
TITLE OF INVENTION: MUTINS OF MAMMALIAN CYTOKINES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,393B
FILING DATE: 01-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0389
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-284-393B-12

Query Match 54.1%; Score 33; DB 1; Length 152;
Best Local Similarity 87.5%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLORPGLQ 8
11111111
Db 12 LTVRPGLO 19

RESULT 10
PCT-US95-08950-12
Sequence 12, Application PC/TUS9508950
GENERAL INFORMATION:
APPLICANT: Zurawski, Sandra M.
APPLICANT: Zurawski, Gerard
TITLE OF INVENTION: MUTINS OF MAMMALIAN CYTOKINES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,393
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0389
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-08950-12

Query Match 54.1%; Score 33; DB 5; Length 152;
Best Local Similarity 87.5%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLORPGLQ 8
11111111
Db 12 LTVRPGLO 19

```
RESULT 11
US-07-952-853-22
; Sequence 22, Application US/07952853
; Patent No. 5863783
GENERAL INFORMATION:
APPLICANT: Van Heuvel, Margaretha
APPLICANT: Bakhuis, Janna G.
APPLICANT: Coustel, Yves
APPLICANT: Harder, Abraham
APPLICANT: De Graaff, Leendert H.
APPLICANT: Filiphi, Michel J. A.
APPLICANT: Van Der Veen, Peter
APPLICANT: Visser, Jacob
APPLICANT: Andreoli, Peter M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
TITLE OF INVENTION: MOLECULES
TITLE OF INVENTION: ENCODING ARABININ-DEGRADING ENZYMES OF FUNGAL
TITLE OF INVENTION: ORIGIN
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/952,853
FILING DATE: 19921125
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 246152003500
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-952-853-22

Query Match          54.1%; Score 33; DB 2; Length 628;
Best Local Similarity 75.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
APPLICANT: Visser, Jacob
APPLICANT: Andreoli, Peter M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
TITLE OF INVENTION: MOLECULES ENCODING ARABININ-DEGRADING ENZYMES OF FUNGAL
TITLE OF INVENTION: ORIGIN
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,848
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,853
FILING DATE: 25-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 246152003500
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-914-848-22

Query Match          54.1%; Score 33; DB 2; Length 628;
Best Local Similarity 75.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 LQORPGLQ 8
      :|||
DB      62 MLONPGLQ 69

RESULT 13
US-08-840-316-3
; Sequence 3; Application US/08840316
; Patent No. 6054567
GENERAL INFORMATION:
APPLICANT: Emerson, Suzanne U., Purcell, Robert H.,
APPLICANT: Tsarev, Sergei, A., and Robinson, Robin A.
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,316
FILING DATE: 11-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4255
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acid residues
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-840-316-3

Query Match 52.5%; Score 32; DB 3; Length 123;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 RRGLOLYFS 12
DB 83 RRGDLVFA 91

RESULT 14
US-08-478-507-9
Sequence 9, Application US/08478507
Patent No. 6120988
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R
APPLICANT: Varbough, Patricia O
APPLICANT: Bradley, Daniel W
APPLICANT: Krawczynski, Krzysztof Z
APPLICANT: Tam, Albert
APPLICANT: Fry, Kirk E
TITLE OF INVENTION: DNA Sequences of Enterically-Transmitted
TITLE OF INVENTION: No. 6120988-A/No. 6120988-B Hepatitis Viral Agent
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,507
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279,823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,078
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921

FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183:22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-507-9

Query Match 52.5%; Score 32; DB 4; Length 123;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 RRGLOLYFS 12
DB 83 RRGDLVFA 91

RESULT 15
US-08-809-523-3
Sequence 3, Application US/08809523
Patent No. 6207416
GENERAL INFORMATION:
APPLICANT: Tsarev, Sergei A., Emerson,
APPLICANT: Suzanne U., Purcell, Robert H.
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,523
FILING DATE: 28-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13102
FILING DATE: 03-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/316,765
FILING DATE: 03-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/947,263
FILING DATE: 18-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459

REFERENCE/DOCKET NUMBER: 2026-4032US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acid residues
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-809-523-3

Query Match 52.5%; Score 32; DB 4; Length 123;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 RPSGLQLYFS 12
||| | |
Db 83 RPSGLDLVFA 91

Search completed: June 4, 2001, 12:14:29
Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:39 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-23

Perfect score: 61

Sequence: 1 LQRPGLQLYFS 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: PIR.67:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	336	2	I48471
2	61	100.0	344	2	A41357
3	61	100.0	374	1	A39878
4	56	91.8	404	2	A46480
5	39	63.9	612	1	A39019
6	38	62.3	393	2	A75079
7	38	62.3	793	2	UC5539
8	38	62.3	1916	2	S30236
9	37	60.7	302	2	C64829
10	37	60.7	307	2	H83705
11	37	60.7	425	2	T32103
12	37	60.7	481	2	S62427
13	37	60.7	805	2	S50277
14	37	60.7	948	2	G83264
15	37	60.7	3386	1	GNWVDF
16	36	59.0	160	2	S77080
17	36	59.0	246	2	D83540
18	36	59.0	315	2	H64120
19	36	59.0	633	2	T17262
20	36	59.0	681	2	T39076
21	36	59.0	766	2	E70916
22	36	59.0	838	1	VGBEGL
23	35	57.4	255	2	T28121
24	35	57.4	257	2	T01254
25	35	57.4	309	2	T31833
26	35	57.4	431	2	T14413
27	35	57.4	498	2	T16190
28	35	57.4	688	2	B65005
29	35	57.4	705	2	S54521

early endosome ant
ribosomal protein
ribosomal protein
ribosomal protein
ribosomal protein
ribosomal protein
progesterone 11alp
aluminum resistanc
3-deoxy-manno-ctu
hypothetical prote
alanine racemase -
alanine racemase,
probable hydrolase
conserved hypothet
tyrosine transamin
hypothetical prote

ALIGNMENTS

RESULT 1
I48471
Fc gamma (IgG) receptor high affinity - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: I48471
R:Prins, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarth, P.M.; Wicker, L.S.; G
Science 260, 695-698, 1993
A>Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for
A:Reference number: I48471; MID:93242399
A:Accession: I48471
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-336 <RES>
A:Cross-references: EMBL:X70980; NID:g311748; PIDN:CAA50311.1; PID:g311749
C:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: Immunoglobulin receptor
F:128-180/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 61; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQRPGLQLYFS 12
DB 231 LQRPGLQLYFS 242

RESULT 2
A41357
Fc gamma (IgG) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C>Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000
C:Accession: A41357; S03019
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A>Title: Isolation and expression of functional high-affinity Fc receptor complementa
A:Reference number: A41357; MID:89100284
A:Accession: A41357
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: GB:X14355; GB:M21090; NID:g31333; PIDN:CAA2536.1; PID:g31334
R:Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A>Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (
A:Reference number: S03018; MID:89098339
A:Accession: S03019
A:Molecule type: mRNA
A:Residues: 1-344 <ALL2>
A:Cross-references: EMBL:X14355; NID:g31333; PIDN:CAA32536.1; PID:g31334

A>Note: the authors translated the codon ACT for residue 25 as Ala
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
 F:117-170/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 61; DB 2; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.00098;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIORPGLQLYFS 12
 |||||
 DB 217 LIORPGLQLYFS 228

RESULT 3

A39878
 Fc gamma (Igc) receptor I-A (high affinity) precursor - human
 N:Alternate names: CD64
 C:Species: Homo sapiens (man)
 C>Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
 C:Accession: A39878; I70304; B41357; S03018; I57525
 R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
 J. Biol. Chem. 266, 13449-13455, 1991
 A>Title: Gene organization of the human high affinity receptor for IgG, Fc gamma RI
 A:Reference number: A39878; MUID:91302383
 A:Accession: A39878
 A:Molecule type: DNA
 A:Status: translated from GB/EMBL/DBJ
 A:Residues: 1-374 <YAN>
 A:Cross-references: GB:M63830; GB:M63835; NID:9180277; PIDN:AAA35678.1; PID:9180279
 R:Porjes, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A>Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: I55577; MUID:93055454
 A:Accession: I70304
 A:Molecule type: mRNA
 A:Status: translated from GB/EMBL/DBJ
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:I03418; NID:9184840; PIDN:AAA36049.1; PID:9292169
 R:Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A>Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: A41357; MUID:89100284
 A:Accession: B41357
 A:Molecule type: mRNA
 A:Residues: 1-24,'S',26-337,'T',339-374 <AL1>
 A:Cross-references: GB:X14356; GB:M21091; NID:931331; PIDN:CAA32537.1; PID:931332
 R:Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A>Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR)
 A:Reference number: S03018; MUID:8998339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24,'S',26-337,'T',339-374 <AL2>
 A:Cross-references: EMBL:X14356; NID:931331; PIDN:CAA32537.1; PID:931332
 R:Perez, C.; Wietzerbin, J.; Benech, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A>Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
 lem.
 A:Reference number: I57525; MUID:93204964
 A:Accession: I57525
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RE2>
 A:Cross-references: GB:S57204; NID:9298692; PIDN:AAD13867.1; PID:94261587
 C:Genetics:
 A:Gene: GDB:FCGR1A; CD64
 A:Cross-references: GDB:I35911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane F
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-292/Domain: extracellular #status predicted <EXT>

F:117-170/Domain: immunoglobulin homology <IMM2>
 F:293-313/Domain: transmembrane #status predicted <TM>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 100.0%; Score 61; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIORPGLQLYFS 12
 |||||
 DB 217 LIORPGLQLYFS 228

RESULT 4

A46480
 Fc gamma (Igc) receptor high affinity - mouse
 N:Alternate names: high affinity IgG receptor
 C:Species: Mus musculus (house mouse)
 C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
 C:Accession: A46480; A43511
 R:Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.
 J. Immunol. 148, 1570-1575, 1992
 A>Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI
 A:Reference number: A46480; MUID:92165399
 A:Accession: A46480
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-404 <OSM>
 A:Note: sequence extracted from NCBI backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212,
 R:Seares, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.
 J. Immunol. 144, 371-378, 1990
 A>Title: Molecular cloning and expression of the mouse high affinity Fc receptor for
 A:Reference number: A43511; MUID:90111035
 A:Accession: A43511
 A:Molecule type: mRNA
 A:Status: preliminary
 A:Residues: 1-404 <SEA>
 A:Cross-references: GB:M1314; NID:9200752; PIDN:AAA40056.1; PID:9200753
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: immunoglobulin receptor; transmembrane protein
 F:127-179/Domain: immunoglobulin homology <IMM>

Query Match 91.8%; Score 56; DB 2; Length 404;
 Best Local Similarity 91.7%; Pred. No. 0.0097;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIORPGLQLYFS 12
 |||||
 DB 226 LIORPGLQLYFS 237

RESULT 5

A39019
 glucose dehydrogenase (acceptor) (EC 1.1.99.10) - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 27-Nov-1991 #sequence_revision 09-May-1997 #text_change 16-Jun-2000
 C:Accession: A39019; S06628
 R:Krasney, P.A.; Carr, C.; Cavenier, D.R.
 Mol. Biol. Evol. 7, 155-177, 1990
 A>Title: Evolution of the glucose dehydrogenase gene in Drosophila.
 A:Reference number: A39019; MUID:90205602
 A:Accession: A39019
 A:Molecule type: DNA
 A:Residues: 1-612 <KRA>
 A:Cross-references: GB:M29298; NID:9157499; PIDN:AAA28571.1; PID:9157500
 R:Whetten, R.; Organ, E.; Krasney, P.; Cox-Foster, D.; Cavenier, D.
 Genetics 120, 475-484, 1988
 A>Title: Molecular structure and transformation of the glucose dehydrogenase gene in
 A:Accession: S06628; MUID:89063537
 A:Molecule type: DNA

A:Residues: 1-96 <MHED>
 A:Cross-references: EMBL:X13582; NID:g8008; PIDN:CAA31918.1; PID:g1841419
 C:Genetics:
 A:Gene: FlyBase:Gld
 A:Cross-references: FlyBase:FBgn0001112
 A:Map position: 3 84C8
 A:Introns: 56/3
 C:Superfamily: alcohol oxidase
 C:Keywords: FAD; flavoprotein; oxidoreductase
 F:66-95/Region: beta-alpha-beta FAD nucleotide-binding fold

Query Match 63.9%; Score 39; DB 1; Length 612;
 Best Local Similarity 77.8%; Pred. No. 21;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 LORPGQLYF 11
 :|||
 Db 415 ERPDQLYF 423

RESULT 6
 A:75079
 hypothetical protein PAB1688 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000
 R:Anonymous; Genoscope
 A:Submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
 A:Reference number: A75001
 A:Accession: A75079
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-393 <RAM>
 A:Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB49934.1; PID:g545844
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB1688
 C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0918

Query Match 62.3%; Score 38; DB 2; Length 393;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LORPGQLYF 11
 :|||
 Db 189 LMRPGLEYF 199

RESULT 7
 A:JC5539
 Smoothed protein precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 17-Mar-1999
 C:Accession: JC5539; PC4476
 R:Akiyama, H.; Shigeno, C.; Hiraki, Y.; Shukunami, C.; Kohno, H.; Akagi, M.; Konishi, J.
 Biochem. Biophys. Res. Commun. 235, 142-147, 1997
 A:Title: Cloning of a mouse Smoothed cDNA and expression patterns of hedgehog signal
 A:Reference number: JC5539; MIMD:97339452
 A:Accession: JC5539
 A:Molecule type: mRNA
 A:Residues: 1-793 <AKR>
 A:Accession: PC4476
 A:Molecule type: protein
 A:Residues: 528-533-539-545-600-605 <AK2>
 A:Experimental source: ADTC5 cell
 C:Comment: This protein is used in the conserved targets in Hedgehog signalling pathway
 g syndromes.
 F:1-32/Domain: signal sequence #status predicted <SIG>

Query Match 62.3%; Score 38; DB 2; Length 793;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 LORPGQLYF 12
 :|||
 Db 583 LORPGQLYF 594

RESULT 8
 A:S30236
 genome polyprotein - zucchini yellow mosaic virus (strain Singapore) (fragment)
 N:Contains: coat protein; nuclear inclusion protein a; RNA-directed RNA polymerase (E
 C:Species: zucchini yellow mosaic virus, ZYMV
 A:Variety: strain Singapore
 C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 17-Nov-2000
 C:Accession: S30236
 R:Wu, M.; Yeong, C.Y.; Lee, S.C.; Wong, S.M.
 Nucleic Acids Res. 21, 1317, 1993
 A:Title: Nucleotide sequence of the 3' half of zucchini yellow mosaic virus (Singapor
 A:Reference number: S30236; MIMD:93219099
 A:Accession: S30236
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-1016 <MDM>
 A:Cross-references: EMBL:X68509; NID:g288233; PIDN:CAA48521.1; PID:g938312
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
 C:Superfamily: tobacco etch virus genome polyprotein
 C:Keywords: coat protein; genome-linked protein; nucleotidyltransferase; phosphoprote
 F:9-61/Product: VPg protein #status predicted <YPG>
 F:62-44/Product: nuclear inclusion protein a #status predicted <NIA>
 F:445-1010/Product: RNA-directed RNA polymerase #status predicted <POL>
 F:1011-1016/Product: coat protein (fragment) #status predicted <COP>
 F:125/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Query Match 62.3%; Score 38; DB 2; Length 1016;
 Best Local Similarity 45.5%; Pred. No. 51;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 LORPGQLYF 11
 :|||
 Db 179 LMRPGQLYF 189

RESULT 9
 A:C64829
 transcription regulator ycan - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 29-Sep-1999
 C:Accession: C64829
 R:Battner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MIMD:97426617
 A:Accession: C64829
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-302 <BIAT>
 A:Cross-references: GB:AE000192; GB:U00096; NID:g1787125; PIDN:AAC73986.1; PID:g17871
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: ycan
 C:Superfamily: hypothetical protein b1328
 C:Keywords: DNA binding; transcription regulation
 F:19-49/Region: helix-turn-helix motif

Query Match 60.7%; Score 37; DB 2; Length 302;
 Best Local Similarity 55.6%; Pred. No. 25;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 ORPGLOLYF 11
 :||| |:
 Db 263 ERGPOIY 271

RESULT 10

ABC transporter (ATP-binding protein) BH0448 [imported] - Bacillus halodurans (strain C-3)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
 C:Accession: H83705
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: AB8650; MIMD:20263314
 A:Accession: H83705
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-307 <STO>
 A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA804167.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH0448
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 60.7%; Score 37; DB 2; Length 307;
 Best Local Similarity 63.6%; Pred. No. 25;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LRPGLOLYF 11
 :||| |:
 Db 274 LRPGLOLYF 284

RESULT 11

hypothetical protein F45C12.14 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T32103
 R:Johnson, D.; Kramer, J.; Keppler, D.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid F45C12.
 A:Reference number: Z21123
 A:Accession: T32103
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-425 <JOH>
 A:Cross-references: EMBL:AF016684; PIDN:AB66212.1; GSPDB:GN00020; CESP:F45C12.14
 C:Genetics:
 A:Gene: CESP:F45C12.14
 A:Map position: 2
 A:Introns: 27/2; 337/3

Query Match 60.7%; Score 37; DB 2; Length 425;
 Best Local Similarity 72.7%; Pred. No. 34;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 LRPGLOLYF 11
 :||| |:
 Db 126 LRPGLOLYF 136

RESULT 12

G-protein signaling regulator fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
 C:Accession: T381179; S62427
 R:Lyte, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, October 1995
 A:Reference number: Z21776
 A:Accession: T381179
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-481 <LY2>
 A:Cross-references: EMBL:Z5485; NID:g1008429; PIDN:CAA91077.1; PID:g1008440; GSPDB:G
 C:Genetics:
 A:Gene: rgs1; SPAC22F3.12c
 A:Map position: 1L

Query Match 60.7%; Score 37; DB 2; Length 481;
 Best Local Similarity 54.5%; Pred. No. 38;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LRPGLOLYF 11
 :||| |:
 Db 348 LRPGLOLYF 358

RESULT 13

ubiquitin-specific proteinase UBP5 (EC 3.4.22.1) - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 29-Sep-1999
 C:Accession: S50277; S50647
 R:Xiao, W.; Fontaine, T.; Tang, M.
 Yeast 10, 1497-1502, 1994
 A:Title: UBP5 encodes a putative yeast ubiquitin-specific protease that is related to
 A:Reference number: S50277; MIMD:95176708
 A:Accession: S50277
 A:Molecule type: DNA
 A:Residues: 1-805 <XIA>
 A:Cross-references: EMBL:U10082; NID:g595376; PIDN:AAC48928.1; PID:g595377
 R:Dietrich, F.S.
 submitted to the EMBL Data Library, December 1994
 A:Description: The sequence of S. cerevisiae cosmid 8229, 9115, 9132, 9981, and lamb
 A:Reference number: S50647
 A:Accession: S50647
 A:Molecule type: DNA
 A:Residues: 1-805 <DIE>
 A:Cross-references: EMBL:U18917; NID:g603377; PIDN:AB64671.1; PID:g603384; MIPS:YER1
 C:Genetics:
 A:Gene: SGD:UBP5
 A:Cross-references: SGD:S0000946; MIPS:YER144c
 A:Map position: 5R
 C:Superfamily: deubiquinating enzyme SSV7
 C:Keywords: hydrolase

Query Match 60.7%; Score 37; DB 2; Length 805;
 Best Local Similarity 63.6%; Pred. No. 63;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LRPGLOLYF 12
 :||| |:
 Db 337 VORPOLPLYS 347

RESULT 14

hypothetical protein PA3064 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: G83264
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: AB2950; MUID:20437337
A:Accession: G83264
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-948 <STO>
A:Cross-references: GB:AE004720; GB:AE004091; NID:g9949154; PIDN:ANG06452.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3064

Query Match 60.7%; Score 37; DB 2; Length 948;
Best Local Similarity 60.0%; Pred. No. 73;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QRPGLQLYFS 12
| |||:|:|:
Db 190 QEPGLKLFN 199

RESULT 15

GNMVDL
genome polypeptide - dengue virus type 4
N:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein 5;
nonstructural protein NS4b
C:Species: dengue virus type 4
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 17-Nov-2000
C:Accession: A94352; A94364; A26897; A29121
R:Zhao, B.; Mackow, E.; Buckler-White, A.; Markoff, L.; Chanock, R.M.; Lal, C.J.; Makino
Virology 155, 77-88, 1986
A>Title: Cloning full-length dengue type 4 viral DNA sequences: analysis of genes coding
A:Reference number: A94352; MUID:87044106
A:Accession: A94352
A:Molecule type: genomic RNA
A:Residues: 1-776 <ZNA>
A:Cross-references: GB:M14931
R:Mackow, E.; Makino, Y.; Zhao, B.; Zhang, Y.M.; Markoff, L.; Buckler-White, A.; Guller,
Virology 159, 217-228, 1987
A>Title: The nucleotide sequence of dengue type 4 virus: analysis of genes coding for nc
A:Reference number: A94364; MUID:87293881
A:Accession: A94364
A:Molecule type: genomic RNA
A:Residues: 774-3386 <MAC>
A:Cross-references: GB:M17255
C:Superfamily: yellow fever virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F:2-113/Product: capsid protein #status predicted <CAP>
F:42-58/Domain: transmembrane #status predicted <TM1>
F:100-116/Domain: transmembrane #status predicted <TM2>
F:114-279/Product: membrane protein precursor #status predicted <MEP>
F:114-204/Domain: nonterminal signal sequence #status predicted <SIG>
F:205-279/Product: membrane protein #status predicted <MEM>
F:267-283/Domain: transmembrane #status predicted <TM3>
F:280-773/Product: envelope protein #status predicted <ENV>
F:728-744/Domain: transmembrane #status predicted <TM4>
F:753-769/Domain: transmembrane #status predicted <TM5>
F:774-1184/Product: nonstructural protein NS1 #status predicted <NS1>
F:1157-1179/Domain: transmembrane #status predicted <TM6>
F:1185-1343/Product: nonstructural protein NS2a #status predicted <N2A>
F:1344-1473/Product: nonstructural protein NS2b #status predicted <N2B>
F:1474-2091/Product: nonstructural protein NS3 #status predicted <NS3>
F:1666-1673/Region: nucleotide-binding motif A (P-loop)
F:1753-1760/Region: DEAH motif
F:2092-2374/Product: nonstructural protein NS4a #status predicted <N4A>
F:2375-2486/Product: nonstructural protein NS5 #status predicted <NS5>
F:2487-3386/Product: nonstructural protein NS5 #status predicted <NS5>
F:182,346,432,730,903,980,2296,2300,2341,2382,2452,2582,2639,2699,2719,2913,3310/Binding

Query Match 60.7%; Score 37; DB 1; Length 3386;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 PGLQLYFS 12
| |||:|:|:
Db 466 PGLQLYFS 473

Search completed: June 4, 2001, 12:15:41
Job time: 274 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:02 ; Search time 37.56 Seconds
(without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-23

Sequence: 1 LLORPGQLVYS 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	374	1 FCGI_HUMAN	P12314 homo sapien
2	56	91.8	404	1 FCGI_MOUSE	P26151 mus musculu
3	40	65.6	3066	1 POLG_BOMVN	O65399 b genome po
4	39	63.9	625	1 DRGL_DROME	P18173 drosophila
5	38	62.3	787	1 SKO_HUMAN	O99835 homo sapien
6	38	62.3	793	1 SMO_MOUSE	P56726 mus musculu
7	38	62.3	793	1 SMO_RAT	P97698 rattus norv
8	38	62.3	3080	1 POLG_ZYMWV	P18479 z genome po
9	38	62.3	3082	1 POLG_ZYMWV	O36879 z genome po
10	38	62.3	3083	1 POLG_ZYMWV	O89330 z genome po
11	37	60.7	302	1 YCAN_ECOLI	P75836 escherichia
12	37	60.7	481	1 YABG_SCHPO	O09777 schizosach
13	37	60.7	805	1 UBP5_YEAST	P39944 saccharomy
14	37	60.7	3386	1 POLG_DENA	P08666 d genome po
15	36	59.0	315	1 PSYC_HAEIN	P45191 haemophilus
16	36	59.0	681	1 YDHE_SCHPO	O92359 schizosach
17	36	59.0	838	1 VGLH_HSV11	P06477 herpes simp
18	36	59.0	838	1 VGLH_HSV11	P08356 herpes simp
19	35	57.4	255	1 YS84_CAEEL	O09381 caenorhabd
20	35	57.4	705	1 YM37_YEAST	O03824 saccharomy
21	34	55.7	116	1 R88_MUSAC	O42862 mus acumi
22	34	55.7	130	1 R88_BUCAC	P46180 buchnera ap
23	34	55.7	133	1 R88_MYCGA	O52346 mycoplasma
24	34	55.7	134	1 R88_ARATH	P56801 arabidopsis
25	34	55.7	134	1 R88_EPIVI	P30058 epifagus vl
26	34	55.7	134	1 R88_TOBAC	P06363 nicotiana t
27	34	55.7	136	1 R88_MAIZE	P08530 zea mays (m
28	34	55.7	136	1 R88_ORISA	P12148 oryza sativ
29	34	55.7	234	1 KDSB_AOUAE	O66161 aquifex aeo
30	34	55.7	377	1 ALR_HELPJ	O92K99 helicobacte
31	34	55.7	377	1 ALR_HELPJ	O25595 helicobacte
32	34	55.7	418	1 Y824_DEIRA	O91445 delnoccocu
33	34	55.7	454	1 ATTY_RAT	P04694 rattus norv

34	34	55.7	763	1 G6PE_RABIT	P56201 oryctolagus
35	34	55.7	843	1 EF2_BERVU	O23755 beta vulgar
36	34	55.7	1270	1 MYPC_MOUSE	O70468 mus musculu
37	34	55.7	1274	1 MYPC_HUMAN	O14896 homo sapien
38	34	55.7	3066	1 POLG_SEMWG	O90059 s genome po
39	33	54.1	129	1 POLG_ECOLI	P02361 escherichia
40	33	54.1	129	1 R88_HAEIN	P44377 haemophilus
41	33	54.1	132	1 R88_GUTH	O46907 guillardia
42	33	54.1	137	1 COMD_HAEIN	P31771 haemophilus
43	33	54.1	152	1 IL3_HUMAN	P08700 homo sapien
44	33	54.1	152	1 IL3_PANTR	O28809 pan troglod
45	33	54.1	310	1 HPRK_ENTFA	O07664 enterococcu

ALIGNMENTS

RESULT	1	STANDARD	PRT	374 AA
FCGI_HUMAN				
ID	P12314; P12315;			
AC	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I) (CD64).			
DE	FCGRI OR FCGRI OR FCGI OR IGFRI.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89098339; PubMed=2974947;			
RA	Allen J.M., Seed B.;			
RT	"Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcRI)."			
RL	Nucleic Acids Res. 16:11824-11824(1988).			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89100284; PubMed=2911749;			
RA	Allen J.M., Seed B.;			
RT	"Isolation and expression of functional high-affinity Fc receptor complementary DNAs."			
RL	Science 243:378-381(1989).			
CC	-1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN, GAMMA. HIGH AFFINITY RECEPTOR.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.			
CC	-1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.			
CC	-1- DATABASE: NAME=PROW; NOTE=CD guide C064 entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/c064.htm".			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL: X14356; CAA32537.1; -			
CC	EMBL: X14355; CAA32536.1; -			
CC	PIR: S03018; S03018.			
CC	PIR: S03019; S03019.			
CC	PIR: A41357; A41357.			
CC	PIR: B41357; B41357.			
CC	HSSP: P12319; 1ALT.			
CC	MIM: 146760; -			
CC	InterPro: IPR003006; -			
CC	Pfam: PF00047; 1g; 3.			

KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain; Alternative splicing; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 374
 FT DOMAIN 22 292
 FT TRANSMEM 293 313
 FT DOMAIN 314 374
 FT DOMAIN 36 92
 FT DOMAIN 117 175
 FT DOMAIN 205 267
 FT DISOLFID 43 85
 FT DISOLFID 124 168
 FT DISOLFID 212 260
 FT CARBOHYD 59 59
 FT CARBOHYD 78 78
 FT CARBOHYD 152 152
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT VARSPLIC 333 374
 FT VARIANT 26 26
 FT VARIANT 59 59
 FT CONFLICT 25 25
 FT SEQUENCE 374 AA: 42605 MW: 2C2A8103ECF16E6 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 61; DB 1; Length 374;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLOREGLQLYFS 12
 DB 217 LLOREGLQLYFS 228

RESULT 2
 FCGL_MOUSE STANDARD: PRT: 404 AA.
 AC P26151;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I).
 DE Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sclirognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90111035; PubMed=2136886;
 RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RT "Molecular cloning and expression of the mouse high affinity FC receptor for IgG.";
 RL J. Immunol. 144:371-378(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92166399; PubMed=1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RT "Structure and mapping of the gene encoding mouse high affinity FC gamma RI and chromosomal location of the human FC gamma RI gene.";
 RL J. Immunol. 148:1570-1575(1992).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: M31314; AAA40056.1;
 CC PIR: A43511; A43511.
 CC PIR: A46480; A46480.
 CC HSSP: P12319; 1ALT.
 CC MGD: MGI:95498; Fcgr1.
 CC InterPro: IPR003006; -.
 CC Pfam: PF00047; 1g; 3.
 CC IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 24
 FT CHAIN 25 404
 FT DOMAIN 25 297
 FT TRANSMEM 298 320
 FT DOMAIN 321 404
 FT DOMAIN 46 66
 FT DOMAIN 67 124
 FT DOMAIN 154 216
 FT CARBOHYD 28 28
 FT CARBOHYD 48 48
 FT CARBOHYD 69 69
 FT CARBOHYD 168 168
 FT CARBOHYD 249 249
 FT SEQUENCE 404 AA: 44887 MW: 1C4F0033842767E7 CRC64;

Query Match
 Best Local Similarity 91.8%; Score 56; DB 1; Length 404;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLOREGLQLYFS 12
 DB 226 LLOREGLQLYFS 237

RESULT 3
 POLG_BCMVN STANDARD: PRT: 3066 AA.
 AC Q65399;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GENOME POLYPROTEIN (CONTAINS: N-TERMINAL PROTEIN (P1); HELPER COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2); GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA) (EC 3.4.22.-) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48); COAT PROTEIN (CP)).
 DE Bean common mosaic virus (strain NL-3 / Michigan) (BCMVN).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=12196;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96191623; PubMed=8607279;
 RA Fang G.W., Allison R.F., Zambolim E.M., Maxwell D.P., Gilbertson R.L.;
 RT "The complete nucleotide sequence and genome organization of bean common mosaic virus (NL3 strain).";
 RL Virus Res. 39:13-23(1995).
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHDID TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT

MAY BE INVOLVED IN REPLICATION.

-1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.

-1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.

-1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL, PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.

-1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.

-1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.

-1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U19287; AAB02170.1; -

DR MEROPS: C04.001; -

DR MEROPS: C06.001; -

DR MEROPS: S30.001; -

DR InterPro: IPR001205; -

DR InterPro: IPR001456; -

DR InterPro: IPR001592; -

DR InterPro: IPR001730; -

DR InterPro: IPR002540; -

DR Pfam: PF00863; Peptidase_C4; 1.

DR Pfam: PF00851; Peptidase_C6; 1.

DR Pfam: PF01577; Poty_pl; 1.

DR Pfam: PF00767; Poty_coat; 1.

DR Pfam: PF00680; RNA_dep_RNA_pol; 1.

DR PRINTS: PR00966; NIAOPTYPTASE.

KW Hydroxylase; Transferase; Thiol protease; RNA-directed RNA polymerase; Coat protein; Polypeptide; Covalent protein-RNA linkage; Helicase;

KM ATP-binding.

FT CHAIN 1 ? N-TERMINAL PROTEIN.

FT CHAIN ? ? HELPER COMPONENT PROTEINASE.

FT CHAIN ? ? PROTEIN P3.

FT CHAIN ? ? 6 KDA PROTEIN 1.

FT CHAIN ? ? CYTOPLASMIC INCLUSION PROTEIN.

FT CHAIN ? ? 6 KDA PROTEIN 2.

FT CHAIN ? ? GENOME-LINKED PROTEIN.

FT CHAIN ? ? NUCLEAR INCLUSION PROTEIN A.

FT CHAIN ? ? NUCLEAR INCLUSION PROTEIN B.

FT CHAIN ? ? COAT PROTEIN.

FT NP_BIND 1258 1265 ATP (POTENTIAL).

SO SEQUENCE 3066 AA; 350385 MW; E358955297FA3F59 CMC64;

Query Match 65.6%; Score 40; DB 1; Length 3066;

Best Local Similarity 63.6%; Pred. No. 30;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LDRPGLQLYF 11

DB 1978 VVARPGLOAYF 1988

RESULT 4

DHGL_DROME STANDARD; PRT; 625 AA.

AC P18173; Q9Y187;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-OCT-2000 (Rel. 40, last sequence update)

DT 01-OCT-2000 (Rel. 40, last annotation update)

DE GLUCOSE DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.10) [CONTAINS: GLD OR CG1152.

GN GLD OR CG1152.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID:7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-90205602; PubMed-2108306;

RA Krasney P.A., Carr C.M., Cavener D.R.;

RT "Evolution of the glucose dehydrogenase gene in Drosophila.";

RL Mol. Biol. Evol. 7:155-177(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson R.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrelle S., Fleischmann W., Foster C., Gabriellian A.E., Garcia N.S., Gelbart W.M., Glasser K., Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C., Jatali M., Kalush F., Kapen G.H., Ke Z., Kenton J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

RN [3]

RP SEQUENCE OF 1-96 FROM N.A.

RX MEDLINE-89065357; PubMed-3143620;

RA Whetten R., Organ E., Krasney P., Cox-Foster D., Cavener D.R.;

RT "Molecular structure and transformation of the glucose dehydrogenase gene in Drosophila melanogaster.";

RL Genetics 120:475-484(1988).

RN [4]

RP SIMILARITY TO YEAST ALCOHOL OXIDASE.

RX MEDLINE-91163320; PubMed-2002763;

RA Cavener D.R., Krasney P.;

RT "Drosophila glucose dehydrogenase and yeast alcohol oxidase are homologous and share N-terminal homology with other flavoenzymes.";

RL Mol. Biol. Evol. 8:144-150(1991).

RN [5]

RP SELENOCYSTEINE.

RA Perlisky S., Merritt K., Cavener D.;

RT "Incorporation of selenocysteine at a UGA codon of GlD.";

RL (In) Annu. Dros. Res. Conf. 39:414C(1998).

CC -1- FUNCTION: ESSENTIAL FOR CUTICULAR MODIFICATION DURING DEVELOPMENT.

CC -1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR -> D-GLUCONO-1,5-LACTONE + REDUCED ACCEPTOR.

CC -1- COFACTOR: FAD.

CC -1- SUBCELLULAR LOCATION: SECRETED AS PART OF THE SEMINAL FLUID
CC TRANSFERRED TO FEMALES.
CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY
CC -1- CAUTION: REF.5 BELIEVES RESIDUE 613 IS A SELENOCYSTEINE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M29298; AAA28571.1; ALT_SEQ.
CC EMBL: AE003672; AAF54038.1; ALT_SEQ.
CC EMBL: X13582; CA31918.1; -
CC PIR: A39019; A39019.
CC PIR: S06628; S06628.
CC FlyBase: FBgn0001112; Gld.
CC InterPro: IPR000172; -
CC Pfam: PF00732; GMC_OXRED.1.
CC PROSITE: PS00623; GMC_OXRED.1; 1.
CC PROSITE: PS00624; GMC_OXRED.2; 1.
CC Oxidoreductase; Flavoprotein; FAD; Signal; Selenium; Selenocysteine.
CC SIGNAL 1 42
CC CHAIN 43 612 GLUCOSE DEHYDROGENASE [ACCEPTOR] SHORT
CC FT 43 612 GLUCOSE DEHYDROGENASE [ACCEPTOR] SHORT
CC FT NP_BIND 66 95 FAD (ADP PART) (PROBABLE)
CC FT ACT_SITE 549 549 POTENTIAL
CC FT SE_CYS 613 613 PROBABLE
CC FT CONFLICT 484 Q -> R (IN REF. 1).
CC SO SEQUENCE 625 AA; 68387 MW; 02664FC3B820BDFC CRC64;

Query Match 63.9%; Score 39; DB 1; Length 625;
Best Local Similarity 77.8%; Pred. No. 9.6;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 ORPGQLYF 11
Db 415 ERPDQLYF 423

RESULT 5
SMO_HUMAN STANDARD; PRT; 787 AA.
AC 099835;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SMOOTHENED HOMOLOG PRECURSOR (SMO).
GN SMOH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic Lung;
RC MEDLINE=97064168; PubMed=8906787;
RA Stone D.M., Hynes M., Armanini M., Swanson T.A., Gu Q., Johnson R.L.,
RA Scott M.P., Pennica D., Goddard A., Phillips H., Noll M., Hooper J.E.,
RA de Sauvage F., Rosenthal A.;
RT "The tumour-suppressor gene patched encodes a candidate receptor for
RT Sonic hedgehog".
RT Nature 384:129-134 (1996).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS BCC LEU-535 AND GLN-562.
RX MEDLINE=98082974; PubMed=9422511;
RX Xie J., Murone M., Luo S.-M., Ryan A., Gu Q., Zhang C., Bonifas J.M.,
RX Lam C.-W., Hynes M., Goddard A., Rosenthal A., Epstein E.H. Jr.,
RA de Sauvage F.J.;

RT "Activating Smoothened mutations in sporadic basal-cell carcinoma".
RL Nature 391:90-92 (1998).
CC -1- FUNCTION: G PROTEIN-COUPLED RECEPTOR THAT PROBABLY ASSOCIATES WITH
CC THE PATCHED PROTEIN (PTCH) TO TRANSDUCE THE HEDGEHOG'S PROTEIN
CC SIGNAL. BINDING OF SONIC HEDGEHOG (SHH) TO ITS RECEPTOR PATCHED IS
CC THOUGHT TO PREVENT NORMAL INHIBITION BY PATCHED OF SMOOTHENED
CC (SMO).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- DISEASE: DEFECTS IN SMO RESULT IN BASAL CELL CARCINOMAS (BCC), A
CC COMMON CANCER IN HUMAN.
CC -1- SIMILARITY: BELONGS TO FAMILY FZ/SMO OF G-PROTEIN COUPLED
CC RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U84401; AAB41788.1; -
CC EMBL: AF114821; AAD17202.1; -
CC EMBL: AF114819; AAD17202.1; JOINED.
CC EMBL: AF114820; AAD17202.1; JOINED.
CC MIM: 601500; -
CC InterPro: IPR000024; -
CC InterPro: IPR000539; -
CC Pfam: PF01534; Frlz2led; 1.
CC Pfam: PF01392; Fz; 1.
CC PRINTS: PR00489; FRI2ZLED.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Developmental protein; Proto-oncogene; Disease mutation.
CC SIGNAL 1 27
CC CHAIN 28 787
CC FT 28 787
CC FT DOMAIN 234 254
CC FT DOMAIN 255 262
CC FT TRANSMEM 263 283
CC FT DOMAIN 284 314
CC FT TRANSMEM 315 335
CC FT DOMAIN 336 358
CC FT TRANSMEM 359 379
CC FT DOMAIN 380 402
CC FT TRANSMEM 403 423
CC FT DOMAIN 424 451
CC FT TRANSMEM 452 472
CC FT DOMAIN 473 524
CC FT TRANSMEM 525 545
CC FT DOMAIN 546 787
CC FT CARBOHYD 35 35
CC FT CARBOHYD 188 188
CC FT CARBOHYD 309 309
CC FT VARIANT 535 535
CC FT VARIANT 562 562
CC FT VARIANT 562 562
CC FT VARIANT 562 562
CC SO SEQUENCE 787 AA; 86396 MW; 8BAC459B34D13F83 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 787;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LURPGLYF 12
Db 579 LRNPGLYF 590

RESULT 6
SMO_MOUSE STANDARD; PRT; 793 AA.
AC 56726;

Db 583 L10NPG0ELSEFS 594

RESULT 8
POLG_ZYMWC STANDARD; PRT; 3080 AA.
ID POLG_ZYMWC
AC P18479; 089334;
DT 01-NOV-1990 (Rel. 16, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER COMPONENT POLYPROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2); GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA) (EC 3.4.22.-) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48); COAT PROTEIN (CP)].
DE Zucchini yellow mosaic virus (strain California) (ZYMV).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae; Potyviruses.
OC NCBI_TaxID=117128;
OX NCBI_TaxID=117128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95146958; PubMed=7844540;
RA Wisler G.C., Purcifull D.E., Hiebert E.;
RT "Characterization of the P1 protein and coding region of the zucchini yellow mosaic virus.";
RL J. Gen. Virol. 76:37-45(1995).
RN [2]
RP SEQUENCE OF 2694-3080 FROM N.A.
RX MEDLINE=90236320; PubMed=2185142;
RA Gal-On A., Antignus Y., Rosner A., Rascach B.;
RT "Nucleotide sequence of the zucchini yellow mosaic virus capsid-encoding gene and its expression in Escherichia coli.";
RL Gene 87:273-277(1990).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- P1M: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- P1M: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC EMBL: J31350; AAA65559.1;
DR EMBL: M35095; AAA48511.1;
DR PIR: JH0103; JH0103.
DR InterPro: IPR001205;
DR InterPro: IPR001456;
DR InterPro: IPR001592;
DR InterPro: IPR001650;
DR InterPro: IPR001730;
DR InterPro: IPR002540;
DR Pfam: PF00863; Peptidase_C4; 1.
DR Pfam: PF00851; Peptidase_C6; 1.
DR Pfam: PF01577; Poly_P1; 1.
DR Pfam: PF00767; Poly_P1; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR PRINTS: PR00966; NIAPOTYPTASE.

KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW ATP-binding.
FT CHAIN 1 304 N-TERMINAL PROTEIN.
FT CHAIN 305 766 HELPER COMPONENT PROTEINASE.
FT CHAIN 767 ? PROTEIN P3.
FT CHAIN ? 1164 6 KDA PROTEIN 1.
FT CHAIN 1165 1798 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1799 1851 6 KDA PROTEIN 2.
FT CHAIN 1852 2041 GENOME-LINKED PROTEIN.
FT CHAIN 2042 2284 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2285 2801 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2802 3080 COAT PROTEIN.
FT NP_BIND 1249 1256 ATP (POTENTIAL).
FT CONFLICT 2694 2695 LE -> ST (IN REF. 2).
FT CONFLICT 2699 2701 IVS -> LRP (IN REF. 2).
FT CONFLICT 2811 2811 A -> T (IN REF. 2).
FT CONFLICT 2834 2834 G -> S (IN REF. 2).
SO SEQUENCE 3080 AA; 350624 MW; 2A1E501DEA69F73 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 3080;
Best Local Similarity 45.5%; Pred. No. 71;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 L10RPGLOLYF 11
Db 1969 IINKPGIOAYF 1979

RESULT 9
POLG_ZYMVS STANDARD; PRT; 3082 AA.
ID POLG_ZYMVS
AC 036979;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2); DE GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA) (EC 3.4.22.-) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48); COAT PROTEIN (CP)].
DE Zucchini yellow mosaic virus (strain Singapore) (ZYMV).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae; Potyviruses.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae; Potyviruses.
OX NCBI_TaxID=117130;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee K.C., Wong S.M.;
RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- P1M: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- P1M: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

```

CC -----
DR EMBL: AF014811; AAB72004.1; -
DR InterPro: IPR001205; -
DR InterPro: IPR001456; -
DR InterPro: IPR001592; -
DR InterPro: IPR001650; -
DR InterPro: IPR001730; -
DR InterPro: IPR002540; -
DR Pfam: PF00863; Peptidase_C4; 1.
DR Pfam: PF00851; Peptidase_C6; 1.
DR Pfam: PF01577; Poly_P1; 1.
DR Pfam: PF00767; Poly_coat; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00271; helicase_C; 1.
DR PRINTS: PR00966; NIAPOTYPTASE.
KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KM ATP-binding.
FT CHAIN 1 307 N-TERMINAL PROTEIN.
FT CHAIN 308 769 HELPER COMPONENT PROTEINASE.
FT CHAIN 770 1115 PROTEIN P3.
FT CHAIN 1116 1167 6 KDA PROTEIN 1.
FT CHAIN 1168 1801 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1802 1854 6 KDA PROTEIN 2.
FT CHAIN 1855 2044 GENOME-LINKED PROTEIN.
FT CHAIN 2045 2287 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2288 2803 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2804 3083 COAT PROTEIN.
FT NP_BIND 1252 1259 ATP (POTENTIAL).
SQ SEQUENCE 3082 AA; 350505 MW; BF5AC062110EF84A CRC64;

Query Match 62.3%; Score 38; DB 1; Length 3082;
Best Local Similarity 45.5%; Pred. No. 71;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLORPGLOLYE 11
DB 1972 IINKPGIOAYF 1982

RESULT 10
POLG_ZYMVR STANDARD; PRT; 3083 AA.
AC 089330;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
DE COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN
DE 1 (K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (K2);
DE GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
DE (EC 3.4.22.-) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION
DE PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.46);
DE COAT PROTEIN (CP)].
OS Zucchini yellow mosaic virus (strain Reunion Island) (ZYMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Polyvirus.
OX NCBI_TaxID=117129;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95146958; PubMed=7844540;
RA Wisler G.C., Purcifull D.E., Hiebert E.;
RT Characterization of the P1 protein and coding region of the zucchini
RT yellow mosaic virus.
RT J. Gen. Virol. 76:37-45(1995).
RL J. Gen. Virol. 76:37-45(1995).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHD
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE

```

```

CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC -----
CC This SWISS-PROF entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U29569; AAB65558.1; -
DR MEROPS: C04.001; -
DR InterPro: IPR001205; -
DR InterPro: IPR001456; -
DR InterPro: IPR001592; -
DR InterPro: IPR001650; -
DR InterPro: IPR001730; -
DR InterPro: IPR002540; -
DR Pfam: PF00863; Peptidase_C4; 1.
DR Pfam: PF00851; Peptidase_C6; 1.
DR Pfam: PF01577; Poly_P1; 1.
DR Pfam: PF00767; Poly_coat; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00271; helicase_C; 1.
DR PRINTS: PR00966; NIAPOTYPTASE.
KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KM ATP-binding.
FT CHAIN 1 307 N-TERMINAL PROTEIN.
FT CHAIN 308 769 HELPER COMPONENT PROTEINASE.
FT CHAIN 770 1115 PROTEIN P3.
FT CHAIN 1116 1167 6 KDA PROTEIN 1.
FT CHAIN 1168 1801 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1802 1854 6 KDA PROTEIN 2.
FT CHAIN 1855 2044 GENOME-LINKED PROTEIN.
FT CHAIN 2045 2287 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2288 2804 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2805 3083 COAT PROTEIN.
FT NP_BIND 1252 1259 ATP (POTENTIAL).
SQ SEQUENCE 3083 AA; 351156 MW; 35B51B455C20C537 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 3083;
Best Local Similarity 45.5%; Pred. No. 72;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLORPGLOLYE 11
DB 1972 IINKPGIOAYF 1982

RESULT 11
YCAN_ECOLI STANDARD; PRT; 302 AA.
AC P75836;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN DMSG-PFLA INTERGENIC REGION.
DE YCAN.
GN Ycan.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;

```

RA Blatner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregory J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN
 RP
 RC
 RX MEDLINE:97061202; PubMed-8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mochimaru K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horinouchi T.,
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC in the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC
 DR EMBL: AE000192; AAC73986.1;
 DR EMBL: D90728; BAA55635.1;
 DR Ecogene; EG13698; ycan.
 DR InterPro; IPR000847;
 DR Pfam; PF00126; HTH_1; 1.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
 DR Hypothetical protein; transcription regulation; DNA-binding.
 KW DNA_BIND 20
 FT H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 302 AA; 34294 MW; 8FF02FB02B24B73 CRC64;
 OY
 DB 263 ERPGQIY 271
 YQ 3 ORPGQLYF 11
 :||| |||
 Query Match 60.7%; Score 37; DB 1; Length 302;
 Best Local Similarity 55.6%; Pred. NO. 11;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

CC the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC
 DR EMBL: Z54285; CAA91077.1;
 DR InterPro; IPR000342;
 DR InterPro; IPR000591;
 DR Pfam; PF00610; DEP; 1.
 DR Pfam; PF00615; RGS; 1.
 DR PRINTS; PRO1301; RGSPTROEIN.
 DR PROSITE; PS50186; DEP; 1.
 DR PROSITE; PS50132; RGS; 1.
 KW Hypothetical protein.
 FT DOMAIN 232 312
 FT DOMAIN 344 474
 FT RGS.
 SQ SEQUENCE 481 AA; 54362 MW; FE80CB5E7BE8BA9D CRC64;

Query Match 60.7%; Score 37; DB 1; Length 481;
 Best Local Similarity 54.5%; Pred. NO. 17;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LIQREQLYF 11
 DB 348 LIKPKMLQTYF 358
 :||| |||
 Query Match 60.7%; Score 37; DB 1; Length 481;
 Best Local Similarity 54.5%; Pred. NO. 17;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 13
 ID UBPS_YEAST STANDARD; PRT; 805 AA.
 AC P39944;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 5 (EC 3.1.2.15) (UBIQUITIN
 DE THIOLESTERASE 5) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 5)
 DE (DEBIOUITINATING ENZYME 5).
 GN UBPS OR YER144C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN
 RP
 RP SEQUENCE FROM N.A.
 RX MEDLINE:95176708; PubMed-7871889;
 RA Xiao W., Fontana T., Tang M.;
 RT "UBP5 encodes a putative yeast ubiquitin-specific protease that is
 RT related to the human Tye-2 oncogene product.";
 RL Yeast 10:1497-1502(1994).
 [2]
 RN
 RP
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan J., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Huntke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Laskari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yellon M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC
 CC -1- CATALYTIC ACTIVITY: UBIQUITIN C-TERMINAL THIOLESTER + H(2)O =
 CC UBIQUITIN + A THIOL.
 CC
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19; ALSO KNOWN AS
 CC FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U10082; AAC48928.1; -
 DR EMBL: U18917; AAB64671.1; -
 DR MEROPS: C19.006; -
 DR SGD: S0000946; UBP5.
 DR InterPro: IPR001394; -
 DR InterPro: IPR001763; -
 DR Pfam: PF00581; Rhodanese; 1.
 DR Pfam: PF00442; UCH-1; 1.
 DR Pfam: PF00443; UCH-2; 1.
 DR PROSITE: PS00972; UCH_2_1; 1.
 DR PROSITE: PS00973; UCH_2_2; 1.
 DR PROSITE: PS00235; UCH_2_3; 1.
 DR Ubiquitin conjugation; Hydrolyase; Thiol protease; Multigene family.
 KW ACT_SITE 455 455 BY SIMILARITY.
 FT ACT_SITE 753 753 BY SIMILARITY.
 FT ACT_SITE 761 761 BY SIMILARITY.
 SO SEQUENCE 805 AA; 92260 MW; 884054A6370DF7 CRC64;
 Query Match 60.7%; Score 37; DB 1; Length 805;
 Best Local Similarity 63.6%; Pred. No. 29;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 LORRGQLYFS 12
 DB 337 VORQPLELYFS 347
 RESULT 14
 ID POLG_DENA STANDARD; PRT: 3386 AA.
 AC P09866; Q88661; Q88662; Q88663; Q88664; Q88665; Q88666; Q88667;
 AC Q88668; Q88669; Q88670; Q88671;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GENOME POLYPROTEIN (CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
 DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;
 DE NONSTRUCTURAL PROTEINS NS1, NS2, NS4A AND NS4B; HELICASE (NS3);
 DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (NS5)].
 OS Dengue virus type 4
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OC NCBI_TaxID=11070;
 OX [1]
 RN SEQUENCE OF 1-776 FROM N.A.
 RP MEDLINE=87044106; PubMed=3022479;
 RA Zhao B., Mackow E., Buckler-White A., Markoff L., Chanock R.M.,
 RA Lai C.-J., Makino Y.;
 RT "Cloning full-length dengue type 4 viral DNA sequences: analysis of
 RT genes coding for structural proteins.";
 RL Virology 155:77-86(1986).
 RN [2]
 RP SEQUENCE OF 774-3386 FROM N.A.
 RX MEDLINE=87293881; PubMed=3039728;
 RA Mackow E., Makino Y., Zhao B., Zhang Y.M., Markoff L.,
 RA Buckler-White A., Guller M., Chanock R., Lai C.J.;
 RT "The nucleotide sequence of dengue type 4 virus: analysis of genes
 RT coding for nonstructural proteins.";
 RL Virology 159:217-228(1987).
 RN [3]
 RP PROCESSING OF THE M PROTEIN.
 RC STRAIN=814669;
 RX MEDLINE=89311624; PubMed=2501515;
 RA Markoff L.;
 RT "In vitro processing of dengue virus structural proteins: cleavage of
 RT the pre-membrane protein.";
 RL J. Virol. 63:3345-3352(1989).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND RNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14931; AAA42964.1; -
 DR PIR: A26897; GNMVDF.
 DR HSSP: P14336; 1SVB.
 DR MEROPS: S07.002; -
 DR InterPro: IPR000069; -
 DR InterPro: IPR000208; -
 DR InterPro: IPR000336; -
 DR InterPro: IPR000404; -
 DR InterPro: IPR000487; -
 DR InterPro: IPR000752; -
 DR InterPro: IPR001123; -
 DR InterPro: IPR001157; -
 DR InterPro: IPR001528; -
 DR InterPro: IPR001850; -
 DR InterPro: IPR002535; -
 DR Pfam: PF01004; Flav1.M; 1.
 DR Pfam: PF00948; Flav1.NS1; 1.
 DR Pfam: PF01005; Flav1.NS2A; 1.
 DR Pfam: PF01002; Flav1.NS2B; 1.
 DR Pfam: PF01350; Flav1.NS4A; 1.
 DR Pfam: PF01349; Flav1.NS4B; 1.
 DR Pfam: PF00972; Flav1.NS5; 1.
 DR Pfam: PF01003; Flav1.Capsid; 1.
 DR Pfam: PF00869; Flav1.glycoprot; 1.
 DR Pfam: PF00949; Flav1.helicase; 1.
 DR Pfam: PF01570; Flav1.propep; 1.
 DR Core protein; glycoprotein; Transferrase; RNA-directed RNA polymerase;
 DR Core protein; coat protein; Envelope protein; Helicase; ATP-binding;
 KW Transmembrane; Nonstructural protein.
 FT CHAIN 1 113
 FT PROPEP 114 204
 FT CHAIN 205 279
 FT CHAIN 280 773
 FT CHAIN 774 1184
 FT CHAIN 1185 1343
 FT CHAIN 1344 1473
 FT CHAIN 1474 2091
 FT CHAIN 2092 2374
 FT CHAIN 2375 2486
 FT CHAIN 2487 3386
 FT DOMAIN 377 390
 FT ACT_SITE 1524 1524
 FT ACT_SITE 1548 1548
 FT ACT_SITE 1608 1608
 FT NP_BIND 1666 1673
 FT SITE 1757 1760
 FT TRANSMEM 42 58
 FT TRANSMEM 267 283
 FT TRANSMEM 728 744
 FT TRANSMEM 753 769
 FT TRANSMEM 1157 1179
 FT DISULFID 282 309
 FT DISULFID 339 395
 FT DISULFID 353 384
 FT DISULFID 371 400
 FT DISULFID 464 564
 FT DISULFID 581 612
 FT CARBOHYD 182 182
 FT CARBOHYD 346 346
 ENVELOPE GLYCOPROTEIN M.
 MAJOR ENVELOPE PROTEIN E.
 NONSTRUCTURAL PROTEIN NS1.
 NONSTRUCTURAL PROTEIN NS2A.
 NONSTRUCTURAL PROTEIN NS2B.
 HELICASE (NS3).
 NONSTRUCTURAL PROTEIN NS4A.
 NONSTRUCTURAL PROTEIN NS4B.
 RNA-DIRECTED RNA POLYMERASE (NS5).
 INVOLVED IN FUSION.
 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ATP (POTENTIAL).
 DEAD BOX.
 POTENTIAL.
 POTENTIAL.
 POTENTIAL.
 POTENTIAL.
 POTENTIAL.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 903 903 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 2236 2236 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 2300 2300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 2452 2452 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 3386 AA; 378903 MW; 5A984B8742C54021 CRC64;
 SQ SEQUENCE

Query Match 60.7%; Score 37; DB 1; Length 3386;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 PGQLYFS 12
 DB 466 PGLELYFN 473

RESULT 15
 PSTC_HAEIN STANDARD; PRT; 315 AA.
 AC P45191;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PHOSPHATE TRANSPORT SYSTEM PERMEASE PROTEIN PSTC.
 GN PSTC OR H11382.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OC NCBI_TaxID-727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE-95350630; PubMed-7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
 Weidman J.F., Phillips C.A., Spriggs T., Heddlom E., Cotton M.D.,
 Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RAE Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT *Whole genome sequencing and assembly of Haemophilus
 RT influenzae Rd.*
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 CC FOR PHOSPHATE; PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE
 CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE CYSTM
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32818; AAC23027.1;
 CC TIGR: H11382;
 DR InterPro: IPR000515;
 DR Pfam: PF00528; BPD.transp. 1.
 DR PROSITE, PS00402; BPD.TRANSP_INN_MEMBR. 1.
 KW Transport; Phosphate transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 22
 FT TRANSMEM 23 43
 FT POTENTIAL.

FT DOMAIN 44 77 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 78 98 POTENTIAL.
 FT DOMAIN 99 117 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 118 138 POTENTIAL.
 FT DOMAIN 139 164 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 165 185 POTENTIAL.
 FT DOMAIN 186 223 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 224 244 POTENTIAL.
 FT TRANSMEM 245 281 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 282 302 POTENTIAL.
 FT DOMAIN 303 315 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 315 AA; 34343 MW; 8FF6FD702A85F0B CRC64;

Query Match 59.0%; Score 36; DB 1; Length 315;
 Best Local Similarity 75.0%; Pred. No. 18;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 PGQLYFS 12
 DB 154 PGLELYFN 161

Search completed: June 4, 2001, 12:24:04
 Job time: 566 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:23:22 ; Search time 107.68 Seconds
(without alignments)
13.062 Million cell updates/sec

Title: US-09-284-107-23
Perfect score: 61
Sequence: 1 LORPGQLYFS 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

SPREMBL_15:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	61	100.0	374	4	092663	092663 homo sapien
2	61	100.0	375	4	092495	092495 homo sapien
3	41	67.2	349	6	09M2T0	09M2T0 bos taurin
4	39	63.9	612	5	094963	094963 drosophila
5	39	63.9	612	5	09V187	09V187 drosophila
6	38	62.3	393	1	09U2X2	09U2X2 pyrococcus
7	38	62.3	1016	14	005912	005912 zucchini ye
8	37	60.7	307	2	09KFN1	09KFN1 bacillus ha
9	37	60.7	425	5	016743	016743 caenorhabd
10	37	60.7	573	2	09K451	09K451 streptomyce
11	37	60.7	917	5	09VY94	09VY94 drosophila
12	36	59.0	64	5	09M107	09M107 drosophila
13	36	59.0	160	2	055990	055990 synchocyst
14	36	59.0	166	14	069075	069075 human hepe
15	36	59.0	171	4	09Y477	09Y477 homo sapien
16	36	59.0	308	2	09L4T8	09L4T8 enterococu
17	36	59.0	379	5	096964	096964 ascaris suu
18	36	59.0	464	6	09MYP4	09MYP4 sus. scrofa
19	36	59.0	517	10	09SFG3	09SFG3 arabidopsis

20	36	59.0	633	4	09UFS7	09UFS7 homo sapien
21	36	59.0	766	2	006817	006817 mycobacteri
22	36	59.0	803	11	008894	008894 mesocricetu
23	36	59.0	803	11	008896	008896 mesocricetu
24	36	59.0	1337	4	09P2E4	09P2E4 homo sapien
25	35.5	58.2	1256	5	09VF50	09VF50 drosophila
26	35	57.4	201	10	09SKR8	09SKR8 arabidopsis
27	35	57.4	257	10	080447	080447 arabidopsis
28	35	57.4	309	5	016416	016416 caenorhabd
29	35	57.4	431	10	023847	023847 brassica ra
30	35	57.4	471	8	09MT31	09MT31 solanum tub
31	35	57.4	498	5	019845	019845 caenorhabd
32	35	57.4	584	10	09SPU1	09SPU1 oryza sativ
33	35	57.4	632	5	044881	044881 caenorhabd
34	35	57.4	688	2	P77182	P77182 escherichia
35	35	57.4	934	2	005478	005478 streptomyce
36	35	57.4	1193	14	09WID3	09WID3 pepper vein
37	35	57.4	1410	4	014221	014221 homo sapien
38	35	57.4	1411	4	015075	015075 homo sapien
39	34	55.7	133	8	09TLU5	09TLU5 cyanidium c
40	34	55.7	134	6	09N2E7	09N2E7 gorilla gor
41	34	55.7	152	3	002116	002116 rhizopus st
42	34	55.7	222	10	09S5X8	09S5X8 lithospermu
43	34	55.7	223	2	09KAP3	09KAP3 bacillus ha
44	34	55.7	241	5	097170	097170 drosophila
45	34	55.7	256	5	017978	017978 caenorhabd

ALIGNMENTS

RESULT 1
ID 092663 PRELIMINARY; PRT; 374 AA.
AC 092663;
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE FC GAMMA RECEPTOR I.
GN AL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=9305454; Pubmed=1430234;
RA Perges A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RT "Novel Fc gamma receptor I family gene products in human mononuclear
cells.";
RL J. Clin. Invest. 90:2102-2109(1992).
DR EMBL: L03418; AAA36049.1;
DR HSSP: P12319; 1ALT;
DR INTERPRO: IPR003006;
DR PFWM: PFW0047; 1g; 3;
DR PRODOM: PD002534; 1;
SQ SEQUENCE 374 AA; 42632 MW; D3D59398CEFA699 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LORPGQLYFS 12
Db 217 LORPGQLYFS 228
RESULT 2
ID 092495 PRELIMINARY; PRT; 375 AA.

```

AC Q92495;
DT 01-FEB-1997 (TREMBLREL. 02, Created)
DT 01-FEB-1997 (TREMBLREL. 02, Last sequence update)
DE 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
DE FC GAMMA RECEPTOR TYPE 1 (FC GAMMA RIB-FC GAMMA RECEPTOR).
GN CD64 OR FC-GAMMA-RIB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Benesh P.D., Sasstry K.N., Iyer R.R., Eichbaum Q.G., Raveh D.,
RA Ezekowitz A.B., Sasstry K., Iyer R.R., Elchbaum Q.G., Raveh D.P.,
RA Ezekowitz R.A.;
RT "Definition of interferon gamma-response elements in a novel human FC
RT gamma receptor gene (Fc gamma R1b) and characterization of the gene
RT structure."
RT J. Exp. Med. 176:1115-1123(1992).
RL EMBL; M91550; AAA58414.1; JOINED.
DR EMBL; M91550; AAA58414.1; JOINED.
DR EMBL; M91551; AAA58414.1; JOINED.
DR EMBL; M91552; AAA58414.1; JOINED.
DR EMBL; M91553; AAA58414.1; JOINED.
DR EMBL; M91554; AAA58414.1; JOINED.
DR EMBL; M91554; AAA58414.1; JOINED.
DR EMBL; S45709; AAD13842.1; JOINED.
DR EMBL; S45707; AAD13842.1; JOINED.
DR EMBL; S45708; AAD13842.1; JOINED.
DR EMBL; S45704; AAD13842.1; JOINED.
DR EMBL; S45705; AAD13842.1; JOINED.
DR HSSP; P12319; IALT.
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 1g; 3.
DR PRODOM; PD002534; -; 1.
SQ SEQUENCE 375 AA; 42881 MW; A84D464CTDDDF91 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIORPGLQLYFS 12
DB 218 LIORPGLQLYFS 229

RESULT 3
Q9M2T0 PRELIMINARY; PRT; 349 AA.
AC Q9M2T0;
DT 01-OCT-2000 (TREMBLREL. 15, Created)
DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
DE FC GAMMA RECEPTOR 1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Li X., Wang A., Zhang G.;
RT "Molecular cloning and identification of full-length cDNA encoding
RT high affinity FC receptor for bovine IgG (Fc gamma RI).";
RL Vet. Immunol. Immunopathol. 75:151-159(2000).
DR EMBL; AF162866; AAF80477.1;
KW Receptor.

```

```

SQ SEQUENCE 349 AA; 39608 MW; D0B77B2EF9408C02 CRC64;

Query Match 67.2%; Score 41; DB 6; Length 349;
Best Local Similarity 70.0%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 ORPGQLYFS 12
DB 219 ERPGQLYFS 228

RESULT 4
Q94963 PRELIMINARY; PRT; 612 AA.
AC Q94963;
DT 01-FEB-1997 (TREMBLREL. 02, Created)
DT 01-FEB-1997 (TREMBLREL. 02, Last sequence update)
DE 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
DE GLUCOSE DEHYDROGENASE.
GN GLD.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAS 54;
RX MEDLINE=97019051; PubMed=8865667;
RA Hamblin M.T., Aguadro C.F.;
RT "High nucleotide sequence variation in a region of low recombination
RT in Drosophila simulans is consistent with the background selection
RT model."
RL Mol. Biol. Evol. 13:1133-1140(1996).
DR EMBL; U63325; AAB48020.1; JOINED.
DR EMBL; U63324; AAB48020.1; JOINED.
DR FLYBASE; FBgn0013875; Dsim\GLD.
DR INTERPRO; IPR000172;
DR PFAM; PF00732; GMC_OXRED_1.
DR PROSITE; PS00623; GMC_OXRED_1; 1.
DR PROSITE; PS00624; GMC_OXRED_2; 1.
DR PRODOM; PD004359; -; 1.
SQ SEQUENCE 612 AA; 66794 MW; 594D084B47EB484C CRC64;

Query Match 63.9%; Score 39; DB 5; Length 612;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 ORPGQLYF 11
DB 415 ERPDQLYF 423

RESULT 5
Q9V187 PRELIMINARY; PRT; 612 AA.
AC Q9V187;
DT 01-MAY-2000 (TREMBLREL. 13, Created)
DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLREL. 14, Last annotation update)
DE GLD PROTEIN.
GN GLD.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;

```

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Agbayeni A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier A.,
 RA Burlis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Moadari C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spirding A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003672; AAF54038.1;
 DR FLYBASE: FBgn0001112; G1d.
 DR INTERPRO: IPR000172;
 DR PFAM: PF00732; GMC_Oxred; 1.
 DR PROSITE: PS00623; GMC_Oxred; 1;
 DR PROSITE: PS00624; GMC_Oxred; 2; 1.
 SO SEQUENCE 612 AA; 66835 MW; BC162CDB4252BCE0 CRC64;

Query Match 63.9%; Score 39; DB 5; Length 612;
 Best Local Similarity 77.8%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 3 ORPGLQLYE 11
 Db 415 ERPDQLYF 423
 RESULT 6
 ID Q902X2 PRELIMINARY; PRT; 393 AA.
 AC Q902X2;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE HYPOTHETICAL 44.9 KDA PROTEIN.
 GN PAB1688.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ORSAY;
 RA Hellig R.;
 RT "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome

RT structure and evolution";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A7248286; CAB49934.1;
 KW Hypothetical protein.
 SO SEQUENCE 393 AA; 44932 MW; 0720744982DEFF19 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 393;
 Best Local Similarity 45.5%; Pred. No. 32;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LLORPGLQLYE 11
 Db 189 LMKRPGIEYF 199

RESULT 7
 ID 005912 PRELIMINARY; PRT; 1016 AA.
 AC 005912;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE GENOME POLYPROTEIN (FRAGMENT).
 OS Zucchini yellow mosaic virus (strain Singapore) (ZYMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=117130;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93219099; PubMed=8464715;
 RA Wu M., Yeung C.Y., Lee S.C., Wong S.M.;
 RT "Nucleotide sequence of the 3' half of zucchini yellow mosaic virus
 RT (Singapore isolate) genome encoding the 4K protein, protease,
 RT polymerase and coat protein.";
 RL Nucleic Acids Res. 21:1317-1317(1993).
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 CC MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A, OR THE N-TERMINAL PART OF
 CC IT, VPg, IS COVALENTLY LINKED TO THE GENOMIC RNA. NI-A ALSO HAS
 CC PROTEOLYTIC ACTIVITY.
 CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: TO OTHER POTYVIRUSES POLYPROTEINS.
 DR EMBL: X68509; CAA48521.1;
 DR INTERPRO: IPR001205;
 DR INTERPRO: IPR001254;
 DR INTERPRO: IPR001730;
 DR PFAM: PF00680; RNA_dep_RNA_pol; 1.
 DR PFAM: PF00683; Peptidase_C4; 1.
 DR PRINTS: PR00966; NIAPOTYPASE.
 KW Hydrolase; Protease; RNA-directed RNA polymerase; Coat protein;
 KW Polyprotein; Covalent protein-RNA linkage.
 FT NON_TER 1
 FT NON_TER 1016
 SO SEQUENCE 1016 AA; 116070 MW; 6887BEACB24E387 CRC64;

Query Match 62.3%; Score 38; DB 14; Length 1016;
 Best Local Similarity 45.5%; Pred. No. 86;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 1 LLORPGLQLYE 11
 Db 179 IINKPGIOAYF 189

RESULT 8
 ID Q9KFN1 PRELIMINARY; PRT; 307 AA.
 AC Q9KFN1;

DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE ABC TRANSPORTER (ATP-BINDING PROTEIN).
GN BH0448.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF001508; BAB04167.1;
KW ATP-binding
SQ SEQUENCE 307 AA; 33924 MW; 05E5A6182D59066 CRC64;

Query Match 60.7%; Score 37; DB 2; Length 307;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LORPGLQLYF 11
DB 274 LLOREGIONNY 284

RESULT 9
O16743 PRELIMINARY; PRT; 425 AA.
AC 016743;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, last annotation update)
DE COSMID F45C12.
GN F45C12.14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smaison N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegrans.";
RI Nature 368:32-38(1994).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RC Johnson D., Kramer J., Keppler D.;
RA Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RL [3]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RC Waterston R.;
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF016684; AAB6212.1;
DR INTERPRO: IPR001810;
DR PFAM: PF00646; F-box; 1.
SQ SEQUENCE 425 AA; 48488 MW; DA2932E3A29F6A7D CRC64;

Query Match 60.7%; Score 37; DB 5; Length 425;
Best Local Similarity 72.7%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 LORPGLQLYF 11
DB 126 LLOREGIONNY 136

RESULT 10
O9K451 PRELIMINARY; PRT; 573 AA.
AC 09K451;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE I.
GN CTAD.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Seeger K.U., Harris D.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL359215; CAB94657.1;
SQ SEQUENCE 573 AA; 64041 MW; 1FC5DC79F6D6DA220 CRC64;

Query Match 60.7%; Score 37; DB 2; Length 573;
Best Local Similarity 77.8%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LORPGLQLYF 10
DB 80 LARPGLOLYF 88

RESULT 11
O9VY94 PRELIMINARY; PRT; 917 AA.
AC 09VY94;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, last annotation update)
DE GRIP91 PROTEIN.
GN GRIP91.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams W.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Abdayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu T., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Teclor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003493; AAF46309.1;
 DR FLYBASE: FBgn0026429; Grlp91.
 DR INTERPRO: IPR000634;
 DR PROSITE: PS00165; DEHYDRATASE_SER_THR; 1.
 SQ SEQUENCE 917 AA; 103706 MW; 6AE88C211D256BB CRC64;

Query Match 60.7%; Score 37; DB 5; Length 917;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LORPGLOLY 10
 DB 579 LERPGLDIT 587

RESULT 12
 ID 09W107 PRELIMINARY; PRT; 64 AA.
 AC 09W107;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE CG18093 PROTEIN.
 GN CG18093
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Abdayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu T., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Teclor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003460; AAF46996.1;
 DR FLYBASE: FBgn0040661; CG18093.
 DR INTERPRO: IPR000804;
 DR PROSITE: PS00989; CLAT_ADAPTOR_S; UNKNOWN.1.
 SQ SEQUENCE 64 AA; 7304 MW; 0A31867ED48B244C CRC64;

Query Match 59.0%; Score 36; DB 5; Length 64;
 Best Local Similarity 54.5%; Pred. No. 11;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LORPGLOLY 11
 DB 50 LVKPGTKLTF 60

RESULT 13
 ID 055990 PRELIMINARY; PRT; 160 AA.
 AC 055990;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 19.2 KDA PROTEIN.
 GN SLD0743
 OS *Synechocystis* sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCC6803;
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugita M., Tabata S.,
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 644 to 924 of the genome."

RL DNA Res. 2:153-166(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97061201; PubMed-8905231;
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugura M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D64005; BAA10772.1;
 DR INTERPRO: IPR002636;
 DR PRAM: PR01724; DUF29; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 160 AA; 19215 MW; CF6E21A1F009DD02 CRC64;

Query Match 59.0%; Score 36; DB 2; Length 160;
 Best Local Similarity 63.6%; Pred. No. 30;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 LIORPGLQLYF 11
 ||:|:|:|
 DB 94 LIESPGLQKXF 104

RESULT 14
 Q69075 PRELIMINARY; PRT; 166 AA.
 ID 069075;
 AC 069075;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DE 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
 DE GLYCOPROTEIN H (FRAGMENT).
 GN GH.
 OS human herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID-10298;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TS026;
 RX MEDLINE-88258458; PubMed-2838568;
 RA Desai P.J., Schaffer P.A., Minson A.C.;
 RT "Excretion of non-infectious virus particles lacking glycoprotein H by
 RT a temperature-sensitive mutant of herpes simplex virus type 1:
 RT evidence that gH is essential for virion infectivity.";
 RL J. Gen. Virol. 69:1147-1156(1988).
 DR EMBL: D00304; BAA00211.1;
 DR NON_TER 1
 FT NON_TER 166
 SQ SEQUENCE 166 AA; 17805 MW; F194344CB6AF9B9 CRC64;

Query Match 59.0%; Score 36; DB 14; Length 166;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LIORPGLQLYFS 12
 ||:|:|:|
 DB 123 LIRSPGRYVYFS 134

RESULT 15
 Q9Y477 PRELIMINARY; PRT; 171 AA.
 ID Q9Y477;
 AC Q9Y477;
 DT 01-NOV-1999 (TRENBLREL. 12, Created)
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
 DT 01-NOV-1999 (TRENBLREL. 12, Last annotation update)

DE D1159A15.1 (FRAGMENT).
 GN D1159A15.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID-9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA McDougall R.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL022575; CAA18605.1;
 FT NON_TER 1
 FT NON_TER 171
 SQ SEQUENCE 171 AA; 18956 MW; B1119293A4015EE1 CRC64;

Query Match 59.0%; Score 36; DB 4; Length 171;
 Best Local Similarity 70.0%; Pred. No. 33;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LIORPGLQLY 10
 ||:|:|:|
 DB 109 LKRPQLEY 118

Search completed: June 4, 2001, 12:23:24
 Job time: 587 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:28 ; Search time 106.14 Seconds
(Without alignments)
6.463 Million cell updates/sec

Title: US-09-284-107-24

Perfect score: 64

Sequence: 1 FYMSKTLRGRL 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq-0401:*

- 1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	12	19	W60556
2	64	100.0	261	20	Y33183
3	64	100.0	344	13	R22549
4	64	100.0	344	17	R91439
5	64	100.0	344	21	Y96183
6	64	100.0	374	13	R20811
7	64	100.0	374	13	R22550
8	64	100.0	374	17	R91438
9	64	100.0	374	17	W00859
10	64	100.0	374	19	W80448
11	64	100.0	374	19	W97833

12	64	100.0	374	19	W97834	Human Fc receptor
13	64	100.0	374	21	Y96134	Human macrophage-s
14	64	100.0	374	21	Y96226	Human high affinity
15	64	100.0	399	21	B43683	Human cancer assoc
16	41	64.1	377	20	W86195	Human Fc receptor
17	40	62.5	410	12	R12428	Hybrid Fc(gamma)R1
18	37	57.8	66	21	Y87523	Conotoxin peptide
19	37	57.8	99	11	R07302	C-terminal sequenc
20	37	57.8	212	21	B01787	Escherichia coli Y
21	37	57.8	242	21	G07162	Arabidopsis thalia
22	37	57.8	242	21	G28967	Arabidopsis thalia
23	37	57.8	242	21	G37166	Arabidopsis thalia
24	37	57.8	242	21	G38490	Arabidopsis thalia
25	37	57.8	245	21	G07161	Arabidopsis thalia
26	37	57.8	245	21	G28966	Arabidopsis thalia
27	37	57.8	245	21	G37165	Arabidopsis thalia
28	37	57.8	245	21	G38489	Arabidopsis thalia
29	37	57.8	300	21	G07160	Arabidopsis thalia
30	37	57.8	300	21	G28965	Arabidopsis thalia
31	37	57.8	300	21	G37164	Arabidopsis thalia
32	37	57.8	300	21	G38488	Arabidopsis thalia
33	37	57.8	1400	15	R44514	MLL amino acid seq
34	37	57.8	1400	17	R92705	MLL CDNA clone 14-
35	37	57.8	1400	21	B29019	Human MLL gene tel
36	37	57.8	3910	14	R38470	ALL-1 protein. Ho
37	37	57.8	3910	16	R66462	ALL-1 (acute lymph
38	36	56.2	259	15	R54218	L. lactis branched
39	36	56.2	250	21	G14209	Arabidopsis thalia
40	36	56.2	352	21	G14208	Arabidopsis thalia
41	36	56.2	378	21	G14207	Arabidopsis thalia
42	35	54.7	61	21	B21431	Cone snail alpha-c
43	35	54.7	61	21	B21433	Cone snail alpha-c
44	35	54.7	64	21	B21429	Cone snail alpha-c
45	35	54.7	847	20	W85603	Hexosaminidase enz

ALIGNMENTS

RESULT	1
ID	W60556
XX	W60556 standard; peptide: 12 AA.
XX	
AC	W60556:
XX	
DT	18-AUG-1998 (first entry)
XX	
DE	Oligopeptide from extracellular domain of CD64.
XX	
KW	Extracellular domain; CD64; identification; antibody;
KW	Immunohistochemical; immunofluorescent analysis; detection;
KW	cell transformation; mutation; anti; oncogene.
XX	
OS	Synthetic.
XX	
PN	W09815833-A1.
XX	
PD	16-APR-1998.
XX	
PF	07-OCT-1997; 97WO-NL00557.
XX	
PR	08-OCT-1996; 96EP-0202791.
XX	
PA	(UYUT-) RIKUSUNIV UTRCHT.
XX	
PI	De Kruif CA, Logtenberg T;
XX	
DR	WPL; 1998-240964/21.
XX	
PT	Identifying peptide(s) binding specifically to protein target - by
PT	expressing on phase surface and testing for binding to immobilised
PT	oligo:peptide derived from the target, useful for, e.g. identifying
PT	specific antibodies

XX Example 1; Page 29; 40pp; English.

CC Synthetic oligopeptides W60537-61 are derived from the extracellular
CC domain of Cb64. They were synthesised on 25 polyethylene rods as
CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
CC affinity binding to the Cb64-covered rods. The specification describes a
CC method for the identification of a peptide able to bind specifically to a
CC target protein. The method comprises displaying the peptide on the
CC surface of a replicable display package, synthesising oligopeptides
CC derived from the target protein on a solid phase, and testing for binding
CC between the peptide and oligopeptides. The method is used to screen large
CC peptide libraries, especially to detect antibodies, or their fragments,
CC that bind to cell markers or that can differentiate between different
CC forms of the same protein, including bispecific antibodies that bind to
CC two non-overlapping epitopes on the same monomeric antigen or two
CC epitopes on different molecules. The genes/oligonucleotides that encode
CC selected peptides can be isolated and used for recombinant production of
CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
CC immunofluorescent analysis, and also to detect cell transformation caused
CC by mutation in (anti)oncogenes.

SO Sequence 12 AA:

Query Match 100.0%; Score 64; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.5e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYWGSKTLGRN 12
DB 1 fywgsktilgrn 12

RESULT 2
Y33183
Y33183 standard; Protein; 261 AA.

AC Y33183;
DT 15-NOV-1999 (first entry)

DE Human sFc-gammaRI protein fragment.

KW Fc receptor; Fc-gammaRIa; human; FCR; model; three-dimension; 3-D;
KW atomic coordinate; bioactive compound design; computer-assisted;
KW drug design; therapy; inhibitor; Fc-gammaR; Fc-epsilonRI; IgG; IgE;
KW tissue damage; hypersensitivity; inflammatory cell recruitment;
KW inflammatory modulator; Fc-gammaRIa; immune function regulation;
KW anti-inflammatory; immunoprotective; sFc-gammaRI.

OS Homo sapiens.

PN WO9940117-A1.

PD 12-AUG-1999.

PF 04-FEB-1999; 99WO-IB00367.

PR 11-SEP-1998; 98US-0099994.

PR 06-FEB-1998; 98US-0073972.

PA (ILEX-) ILEXUS PTY LTD.

PI Baeil JB, Epa V, Garrett TPJ, Hogarth PM, Matthews BR;

PI Maxwell KF, McCarthy TD, McKenzie IFC, Pieterse GA;

PI Powell MS;

DR WPI; 1999-539978/45.

PT Three-dimensional structures and models of Fc receptors, useful in
XX computer-assisted drug design

PS Claim 11; Page 316-317; 326pp; English.

CC This invention describes a novel model of an Fc receptor (FcR) protein
CC representing a three-dimensional (3-D) structure that substantially
CC conforms to the specified atomic coordinates. Computer model images of
CC the FcR can be used to design bioactive chemical compounds, e.g.
CC oligonucleotides, peptides, peptidomimetics and small organic molecules,
CC by computer-assisted methods of drug design. Therapeutic compositions
CC that inhibit the activity of Fc-gammaR or Fc-epsilonRI can be used to
CC reduce IgG- or IgE-mediated tissue damage, e.g. IgG- or IgE-mediated
CC hypersensitivity, recruitment of inflammatory cells or release of
CC inflammatory mediators. The therapeutic compositions can also be used to
CC treat disease in animals. The knowledge of the 3-D structure of Fc-gamma
CC R1a and models of other FcR provides a means for designing and producing
CC compounds that regulate immune function and inflammation in an animal,
CC including humans (i.e. structure based drug design). For example,
CC chemical compounds can be designed to block binding of immunoglobulin to
CC an Fc receptor protein using various computer programs and models. The
CC products of the invention have anti-inflammatory and immunoprotective
CC activity. This sequence represents the human sFc-gammaRI protein
CC fragment described in the method of the invention.

SO Sequence 261 AA:

Query Match 100.0%; Score 64; DB 20; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYWGSKTLGRN 12
DB 210 fywgsktilgrn 221

RESULT 3
R22549
R22549 standard; Protein; 344 AA.

AC R22549;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FcRI receptor encoded by clone p98.

KW Rapid immunoselection cloning technique; cell surface antigen;
KW immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN WO9201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GENO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amlot M;

PI WPI; 1992-056864/07.

PI N-PSDB; Q21179.

PT New CD53 cell surface antigen and DNA encoding it - for
XX immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

CC This amino acid sequence was predicted from the cDNA sequence of
CC cDNA clone p98/X2. It differs from the sequence predicted from
CC clone p135 (see Q21179) at position 25; p135 encodes a Ser residue
CC while clone p98/X2 predicts a Thr residue. At position 58, p90

CC of the p98 product is truncated compared with those of the p135

FT	/note= "encoded by AAT"
Misc-4:44xxxxxx	313

FT FT /note= "encoded by CAA"
FT Misc-difference 216
FT /note= "encoded by TTC"
FT Misc-difference 220
FT /note= "encoded by CGT"
FT Misc-difference 268
FT /note= "encoded by AAT"
FT Misc-difference 305
FT /note= "encoded by GTG"
FT Misc-difference 306
FT /note= "encoded by AAC"
FT Misc-difference 332
FT /note= "encoded by GGT"
FT Misc-difference 333
FT /note= "encoded by GGC"
FT Misc-difference 338
FT /note= "encoded by CCT"
FT US6111093-A.
XX PD 29-AUG-2000.
XX PF 28-OCT-1998; 98US-0181612.
XX PR 01-DEC-1992; 92US-0983647;
PR 25-FEB-1988; 88US-0160416,
PR 13-JUL-1989; 89US-0379076,
PR 23-MAR-1990; 90US-0498809,
PR 13-JUL-1990; 90US-0535759.
XX (GENO) GEN HOSPITAL CORP.
PA Stamenkovic I, Seed B;
XX WPI: 2000-586382/55.
DR N-PSDB: A50631.
XX

Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
useful for immunodiagnosis and immunotherapy of immune-mediated
infections or disorders, e.g. asthma, immune-complex disease, parasitic
diseases -

Example 10; Column 53-55; 75pp; English.

The present sequence is that of a human macrophage specific FCRI,
as deduced from cDNA clone p135 (see A50631), which was isolated
from a cDNA library expressed in COS cells using a novel method of
the invention designed to isolate CSA nucleic acids. The method is
based upon transient expression of a CSA in eukaryotic cells and
physical selection of cells expressing the antigen by adhesion to
(panning on) an antibody-coated substrate such as a culture dish.
CSA nucleic acids isolated by the method of the invention, and the
proteins they encode, are useful for immunodiagnostic and
immunotherapeutic applications, including the diagnosis and
treatment of immune-mediated infections, diseases, and disorders in
animals, including humans. These disorders include asthma,
immune-complex disease, amyloidosis, parasitic diseases or multiple
sclerosis. FCRI is a high affinity receptor for the Fc portion of
IgG, normally located on the cell surfaces of macrophages. The
ability to interfere with such bonding, or to cause it to occur on
cell surfaces other than macrophages, is useful in therapy. A fusion
protein of FCRI and a receptor ligand will be helpful to increase
the potencies of antibodies in therapy.

Sequence 344 AA;

```

Db      229 fymgsktlirgn 240

RESULT   6
R20811 ID R20811 standard; Protein; 374 AA.
XX AC R20811;
XX DT 21-MAY-1992 (first entry)
DE Human macrophage-specific FcRI receptor encoded by clone p135.
XX
XX Rapid immunoselection cloning technique; cell surface antigen;
KW immunodiagnosis; high affinity receptor.
XX
OS Homo sapiens.
PN WO9201049-A.
PD 23-JAN-1992.
PF 15-JUL-1990; 90WO-US04986.
PR 13-JUL-1990; 90US-0553759.
PA (GEHO-) GEN HOSPITAL CORP.
PI Seed B, Aruffo A., Amlot M;
PS WPL: 1992-056864/07.
DR N-PDB: Q21178.
PT New CD53 cell surface antigen and DNA encoding it - for
PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
XX Example 10; Page 94a; 160pp; English.
CC This amino acid sequence was predicted from the cDNA sequence of
CC cDNA clone p135. It differs from the sequence predicted from
CC clones p98/X2 and p90 (see Q21179 and Q21180, respectively) at
CC position 25; p135 encodes a Ser residue and the other two clones
CC predict a Thr residue. At position 58, p135 predicts Leu and p90
CC predicts Val. Sequences predicted from all 3 clones show the
CC typical features of a type I integral membrane protein and include
CC a short hydrophobic signal sequence, a single 21-residue
CC hydrophobic membrane-spanning domain, and a short, highly charged
CC cytoplasmic domain. The extracellular portion contains six
CC potential N-linked glycosylation sites and six Cys residues
CC distributed among three C2 set Ig-related domains. A fusion protein
CC of FcRI and a receptor ligand will be helpful to increase the
CC potency of antibodies in therapy.
XX
XX Sequence 374 AA;
SQ
Query Match 100.0%; Score 64; DB 13; Length 374;
Best Local Similarity 100.0%; Pred. NO. 0.00028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 1 FYMGSKTLIRGN 12
| | | | | | | | | |
Db 229 fymgsktlirgn 240

RESULT   7
R22550 ID R22550 standard; Protein; 374 AA.
XX AC R22550;
XX DT 21-MAY-1992 (first entry)

```


PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 PI WPI: 1996-200279/20.
 DR N-PSDB; T14719.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 of diagnostic and therapeutic proteins
 XX
 PS Example 10; Column 55-56; 79pp; English.
 XX
 CC The amino acid sequence (W00859) of human FCRI was detd. from a
 CC cDNA clone, p90. (T14719), obt'd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Other cDNA clones (see also T14717-18) coded for
 CC variants (R91438-39) of the FCRI sequence. A fusion protein of
 CC FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.
 CC
 XX
 SQ Sequence 374 AA.

Query Match 100.0%; Score 64; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.00028;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYWGSKTLRGNN 12
 |||||
 DB 229 fywgsktlrgnn 240

RESULT 10
 ID W80448 standard; Protein; 374 AA.
 XX
 AC W80448;
 XX
 DT 07-JUN-1999 (first entry)
 XX
 DE Human Fc receptor I.
 XX
 DE Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 KM cloning.
 KW
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2 /note= "encoded by TGG"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by GAC"
 FT Misc-difference 60 /note= "encoded by CTG"
 FT Misc-difference 77 /note= "encoded by CCC"
 FT Misc-difference 85 /note= "encoded by AAT"
 FT Misc-difference 99 /note= "encoded by TCC"
 FT Misc-difference 103 /note= "encoded by CAA"
 FT Misc-difference 141 /note= "encoded by CCC"
 FT Misc-difference 141 /note= "encoded by GGC"
 FT Misc-difference 159

FT /note= "encoded by AAC"
 FT Misc-difference 171 /note= "encoded by ARG"
 FT Misc-difference 176 /note= "encoded by GTC"
 FT Misc-difference 256 /note= "encoded by GGC"
 XX
 PN US5830731-A.
 XX
 PD 03-NOV-1998.
 XX
 PF 21-MAY-1997; 97US-0861205.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1986; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 21-MAY-1997; 97US-0861205.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 PI WPI: 1998-609251/51.
 DR N-PSDB; V63456.
 XX
 PT New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNA(s), especially
 PT human lymphocyte antigenic sequences
 XX
 PS Example 10; Column 53-54; 75pp; English.
 XX
 CC This is the amino acid sequence of human Fc receptor I (FCRI), as
 CC deduced from cDNA clone p135 (see V63456) isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from polyA RNA of cells
 CC from a single patient undergoing extracorporeal interleukin-2
 CC induction therapy. Clones p90 (see X07372) and p98 (see X07373)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 CC
 XX
 SQ Sequence 374 AA.

Query Match 100.0%; Score 64; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.00028;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYWGSKTLRGNN 12
 |||||
 DB 229 fywgsktlrgnn 240

RESULT 11
 ID W97833 standard; Protein; 374 AA.
 XX
 AC W97833;

XX 07-JUN-1999 (first entry)
 XX Human Fc receptor I.
 XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 XX cloning.
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 2 /note- "encoded by TGG"
 FT Misc-difference 23 /note- "encoded by CTC"
 FT Misc-difference 44 /note- "encoded by GAC"
 FT Misc-difference 45 /note- "encoded by CTC"
 FT Misc-difference 60 /note- "encoded by CCC"
 FT Misc-difference 77 /note- "encoded by AAT"
 FT Misc-difference 85 /note- "encoded by TCC"
 FT Misc-difference 99 /note- "encoded by CAA"
 FT Misc-difference 103 /note- "encoded by CCC"
 FT Misc-difference 141 /note- "encoded by GGC"
 FT Misc-difference 159 /note- "encoded by AAC"
 FT Misc-difference 171 /note- "encoded by ATG"
 FT Misc-difference 176 /note- "encoded by GTC"
 FT Misc-difference 256 /note- "encoded by GGC"
 FT /note- "encoded by GGC"
 XX
 XX US5830731-A.
 XX
 XX 03-NOV-1998.
 XX
 XX 21-MAY-1997; 9705-0861205.
 XX
 XX 01-DEC-1992; 92US-0983647.
 XX 25-FEB-1988; 88US-0160416.
 XX 13-JUL-1989; 89US-0379076.
 XX 13-JUL-1990; 90US-0533759.
 XX 21-MAY-1997; 9705-0861205.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 XX
 XX Aruffo A, Seed B;
 XX
 XX MPI; 1998-609251/51.
 XX N-PSDB; X07372.
 XX
 XX New cloning vector and polylinker - based on existing sequences for
 XX efficient cloning and expression of mammalian cDNA(s), especially
 XX human lymphocyte antigenic sequences
 XX
 XX Example 10; Column 53-54; 75pp; English.
 XX
 XX This is the amino acid sequence of human Fc receptor I (FcRI), as
 XX deduced from cDNA clone p90 (see X07372) isolated using a rapid
 XX immunoselection cloning method from a cDNA library expressed in COS
 XX cells. The cDNA library was constructed from polyA RNA of cells
 XX from a single patient undergoing extracorporeal interleukin-2
 XX induction therapy. Clones p135 (see V63456) and p98 (see X07373)
 XX were also obtained. A novel method for cloning cDNAs from mammalian
 XX expression libraries is based on transient expression of an antigen

CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 XX
 XX Sequence 374 AA;
 XX
 XX
 XX Query Match 100.0%; Score 64; DB 19; Length 374;
 XX Best Local Similarity 100.0%; Pred. No. 0.00028;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 FYMSKTLRGNN 12
 XX |||||
 XX DB 229 fymsktilrgnn 240
 XX
 XX RESULT 12
 XX ID W97834 standard; Protein; 374 AA.
 XX AC W97834;
 XX
 XX 07-JUN-1999 (first entry)
 XX
 XX Human Fc receptor I.
 XX
 XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 XX cloning.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 2 /note- "encoded by TGG"
 FT Misc-difference 23 /note- "encoded by CTC"
 FT Misc-difference 44 /note- "encoded by GAC"
 FT Misc-difference 45 /note- "encoded by CTC"
 FT Misc-difference 60 /note- "encoded by CCC"
 FT Misc-difference 77 /note- "encoded by AAT"
 FT Misc-difference 85 /note- "encoded by TCC"
 FT Misc-difference 99 /note- "encoded by CAA"
 FT Misc-difference 103 /note- "encoded by CCC"
 FT Misc-difference 141 /note- "encoded by GGC"
 FT Misc-difference 159 /note- "encoded by AAC"
 FT Misc-difference 171 /note- "encoded by ATG"
 FT Misc-difference 176 /note- "encoded by GTC"
 FT Misc-difference 256 /note- "encoded by GGC"
 FT /note- "encoded by GGC"
 XX
 XX US5830731-A.

XX 03-NOV-1998.
 PD
 XX
 PF 21-MAY-1997; 97US-0861205.
 XX
 XX 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 21-MAY-1997; 97US-0861205.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PI Aruffo A, Seed B;
 XX
 PI WPI: 1998-609251/51.
 DR N-PSDB; X07373.
 XX
 XX New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNA(s), especially
 PT human lymphocyte antigenic sequences
 XX
 PS Example 10; Column 53-54; 75pp; English.

XX This is the amino acid sequence of human Fc receptor I (FcRI), as
 CC deduced from cDNA clone p98/X2 (see X07373) isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from poly A RNA of cells
 CC from a single patient undergoing extracorporeal Interleukin-2
 CC induction therapy. Clones p135 (see V63456) and p90 (see X07372)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 CC
 SO Sequence 374 AA.

Query Match 100.0%; Score 64; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.00028;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FWMGSKTLGRN 12
 |||||||||
 Db 229 fymgsktlgrn 240

RESULT 13
 Y96134
 ID Y96134 standard; Protein; 374 AA.
 AC Y96134;
 XX
 XX 19-DEC-2000 (first entry)
 DT
 DE Human macrophage-specific FcRI.
 XX
 XX

Macrophage; FcRI; cell surface antigen; human; immunoselection;
 KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

XX Homo sapiens.
 OS
 XX
 FH key
 FT Misc-difference 15
 FT /note= "encoded by GCG"
 FT Misc-difference 38
 FT /note= "encoded by ACC"
 FT Misc-difference 50
 FT /note= "encoded by CCG"
 FT Misc-difference 51
 FT /note= "encoded by ACC"
 FT Misc-difference 55
 FT /note= "encoded by CAC"
 FT Misc-difference 56
 FT /note= "encoded by TTC"
 FT Misc-difference 60
 FT /note= "encoded by CCC"
 FT Misc-difference 64
 FT /note= "encoded by CAG"
 FT Misc-difference 82
 FT /note= "encoded by CAA"
 FT Misc-difference 116
 FT /note= "encoded by GAA"
 FT Misc-difference 117
 FT /note= "encoded by GCA"
 FT Misc-difference 122
 FT /note= "encoded by TTC"
 FT Misc-difference 123
 FT /note= "encoded by ACC"
 FT Misc-difference 126
 FT /note= "encoded by CCG"
 FT Misc-difference 129
 FT /note= "encoded by CAT"
 FT Misc-difference 134
 FT /note= "encoded by AAT"
 FT Misc-difference 136
 FT /note= "encoded by GTT"
 FT Misc-difference 139
 FT /note= "encoded by CCA"
 FT Misc-difference 140
 FT /note= "encoded by AAT"
 FT Misc-difference 213
 FT /note= "encoded by CAA"
 FT Misc-difference 216
 FT /note= "encoded by TTC"
 FT Misc-difference 220
 FT /note= "encoded by CGT"
 FT Misc-difference 268
 FT /note= "encoded by AAT"
 FT Misc-difference 305
 FT /note= "encoded by GCG"
 FT Misc-difference 306
 FT /note= "encoded by AAC"
 FT Misc-difference 332
 FT /note= "encoded by GGT"

US6111093-A.
 PD 29-AUG-2000.
 XX
 XX 28-OCT-1998; 98US-0181612.
 XX
 XX 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 23-MAR-1990; 90US-0498809.
 PR 13-JUL-1990; 90US-0553759.
 XX

(GEHO) GEN HOSPITAL CORP.
 PI Stamenkovic I, Seed B;
 XX

DR WPI: 2000-586382/55.
 DR N-PSDB: A50592.
 XX
 PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 PT useful for immunodiagnosis and immunotherapy of immune-mediated
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 PT diseases
 XX
 PS Example 10; Column 53-55; 75pp; English.
 XX
 CC The present sequence is that of a human macrophage specific FCRI,
 CC as deduced from cDNA clone p135 (see A50592), which was isolated
 CC from a cDNA library expressed in COS cells using a novel method of
 CC the invention designed to isolate CSA nucleic acids. The method is
 CC based upon transient expression of a CSA in eukaryotic cells and
 CC physical selection of cells expressing the antigen by adhesion to
 CC (panning on) an antibody-coated substrate such as a culture dish.
 CC CSA nucleic acids isolated by the method of the invention, and the
 CC proteins they encode, are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders in
 CC animals, including humans. These disorders include asthma,
 CC immune-complex disease, amyloidosis, parasitic diseases or multiple
 CC sclerosis. FCRI is a high affinity receptor for the FC portion of
 CC IgG, normally located on the cell surfaces of macrophages. The
 CC ability to interfere with such bonding, or to cause it to occur on
 CC surfaces other than macrophages, is useful in therapy. A fusion
 CC protein of FCRI and a receptor ligand will be helpful to increase
 CC the potencies of antibodies in therapy.
 CC
 CC Sequence 374 AA:
 SO
 Query Match 100.0%; Score 64; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.00028;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FYMGSKTLRGGRN 12
 DB 229 fymgsktlrggrn 240
 |||||
 RESULT 14
 ID Y96226 standard; Protein; 374 AA.
 AC Y96226;
 XX
 DT 11-SEP-2000 (first entry)
 XX
 DE Human high affinity Fc receptor; Fc gammaRI.
 XX
 DE Human high affinity Fc receptor; Fc gammaRI; immunoglobulin;
 XX
 DE Infection; immune response; CD64; monocyte; macrophage; neutrophil;
 KW eosinophil; HIV; IgG; immunosuppressive; antirheumatic; cytostatic;
 KW antitumour; AIDS; SLE; multiple myeloma; rheumatoid arthritis;
 KW systemic lupus erythematosus; tumour.
 XX
 OS Homo sapiens.
 OS
 PN EP1006183-A1.
 XX
 PD 07-JUN-2000.
 XX
 PF 03-DEC-1998; 98EP-0122969.
 XX
 PR 03-DEC-1998; 98EP-0122969.
 XX
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX
 DR WPI: 2000-367968/32.
 DR N-PSDB: A27466.
 XX

PT Novel Fc receptor lacking transmembrane domains, a signal peptide, and
 PT glycosylation, useful for diagnosing and treating immune disorders and
 PT cancer
 XX
 PS Disclosure; Page 26-28; 60pp; English.
 XX
 CC The present sequence is the human high affinity Fc receptor, Fc gammaRI.
 CC Fc gammaRI is also known as CD64. Fc receptors play an important
 CC role in defending the body against infections. First, pathogens are
 CC opsonised by serum immunoglobulins. The resulting complex then binds to
 CC cells expressing Fc receptors. Fc gammaRI molecules are expressed by
 CC monocytes and macrophages, but expression can also be induced on
 CC neutrophils and eosinophils. Upon Fc receptor activation, immune effector
 CC pathways are activated, leading to immune response. The present sequence
 CC may be modified to produce recombinant versions. The recombinant Fc
 CC receptor consist only of the extracellular portion of the receptor and
 CC are not glycosylated i.e. they do not have transmembrane domains or
 CC signal peptides. The recombinant proteins may be used in immunoassays to
 CC determine the immune status of patients with chronic diseases of the
 CC immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple
 CC myeloma (MM), or rheumatoid arthritis. In addition, pharmaceutical
 CC compositions containing recombinant proteins may be used to treat or
 CC prevent autoimmune diseases, allergies or tumours, especially AIDS,
 CC rheumatoid arthritis or MM.
 CC
 CC Sequence 374 AA:
 SO
 Query Match 100.0%; Score 64; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.00028;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FYMGSKTLRGGRN 12
 DB 229 fymgsktlrggrn 240
 |||||
 RESULT 15
 ID B43683 standard; Protein; 399 AA.
 AC B43683;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1128.
 XX
 DE Human: cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
 KW antidiabetic; antisthmatic; antirheumatic; antiarthritic; antiviral;
 KW antitumour; antihypertensive; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
 KW vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 OS
 PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2000-587533/55.
 DR

DR N-PSDB; C77892.

XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.XX
PS Claim 11; Page 1739-1740; 2352pp; English.

XX C77607 to C78448 encode the human cancer associated proteins given in
CC B43398 to B44239. The proteins can have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic;
CC antiasthmatic; antirheumatic; antiarthritic; antiinflammatory;
CC antithyroid; antiallergic; antibacterial; antiviral; dermatological;
CC neuroprotective; cardiant; thrombolytic; coagulant; neotropic;
CC vasotropic; antipsoriatic and antilangiogenic. The polynucleotides and
CC polypeptides can be used for preventing, treating or ameliorating medical
CC conditions and diagnosing pathological conditions. Polynucleotides,
CC polypeptides, antibodies, agonists and antagonists from the present
CC invention may be used to treat immune disorders by activating or
CC inhibiting the proliferation, differentiation or mobilisation of immune
CC cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC allergic reactions, graft versus host disease and organ rejection,
CC modulate haemostatic or thrombolytic activity, modulate inflammation,
CC cancers, cardiovascular disorders, neurological disease and bacterial or
CC viral infections. The peptides, nucleotides, antibodies, agonists and
CC antagonists may be also be used in drug screens. C78449 to C78457 and
CC B44240 represent sequences used in the exemplification of the present
CC invention.

XX
SQ Sequence 399 AA;

Query Match 100.0%; Score 64; DB 21; Length 399;

Best Local Similarity 100.0%; Pred No. 0.0003;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYWGSRTLRGRN 12

Db 254 fywgsrtlgrn 265

Search completed: June 4, 2001, 12:13:29
Job time: 202 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:29 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-24
Perfect score: 64
Sequence: 1 FYMSKTLRGHN 12

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6ackfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	62.5	91	2	US-08-332-562A-94
2	37	57.8	1400	1	US-08-080-255-7
3	37	57.8	1400	4	US-08-465-713-7
4	37	57.8	1400	5	PCT-US83-05857-7
5	37	57.8	3969	4	US-08-061-376-5
6	36	56.2	259	1	US-08-403-866-5
7	33	51.6	135	5	US-08-469-667-7
8	33	51.6	135	5	PCT-US95-07289-7
9	33	51.6	322	2	US-08-428-414A-2
10	32	50.0	95	2	US-08-341-843B-4
11	32	50.0	95	2	US-08-427-497E-9
12	32	50.0	140	4	US-09-405-035-1
13	32	50.0	140	4	US-09-405-035-2
14	32	50.0	140	4	US-09-405-035-3
15	32	50.0	140	4	US-09-405-035-4
16	32	50.0	140	5	PCT-US94-09789-2
17	32	50.0	306	2	US-08-560-098A-45
18	32	50.0	331	2	US-08-560-098A-46
19	32	50.0	365	1	US-08-093-741-83
20	32	50.0	365	1	US-08-720-012-83
21	32	50.0	393	2	US-08-560-098A-44
22	32	50.0	393	4	US-08-560-098A-44
23	32	50.0	393	4	US-08-867-024C-24
24	32	50.0	411	1	US-08-087-163-1
25	32	50.0	411	1	US-08-286-748B-18
26	32	50.0	411	1	US-08-153-799-18
27	32	50.0	411	2	US-08-560-098A-48

28	32	50.0	430	1	US-07-942-157A-3	Sequence 3, Appl 1
29	32	50.0	430	6	5219568-2	Patent No. 5219568
30	32	50.0	431	6	5188829-1	Patent No. 5188829
31	32	50.0	432	2	US-08-560-098A-47	Sequence 2, Appl 1
32	32	50.0	561	1	US-08-358-901-2	Sequence 2, Appl 1
33	32	50.0	561	1	US-08-566-347-2	Sequence 2, Appl 1
34	32	50.0	561	1	US-08-693-835-2	Sequence 2, Appl 1
35	32	50.0	570	2	US-08-867-364-1	Sequence 2, Appl 1
36	32	50.0	570	2	US-08-867-364-7	Sequence 7, Appl 1
37	32	50.0	570	3	US-09-368-408-1	Sequence 7, Appl 1
38	32	50.0	570	3	US-09-368-408-7	Sequence 7, Appl 1
39	31	48.4	70	3	US-09-003-708A-4	Sequence 4, Appl 1
40	31	48.4	260	3	US-08-906-769-139	Sequence 139, App
41	31	48.4	260	4	US-08-906-616-139	Sequence 139, App
42	31	48.4	260	4	US-08-639-075A-139	Sequence 139, App
43	31	48.4	260	4	US-09-012-431-139	Sequence 31, Appl 1
44	31	48.4	383	1	US-08-464-523B-31	Sequence 60, Appl 1
45	31	48.4	474	4	US-08-961-083-60	

ALIGNMENTS

RESULT 1
US-08-332-562A-94
Sequence 94, Application US/08332562A
Patent No. 5985599
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
TITLE OF INVENTION: PC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-Oct-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-94
Query Match 62.5% ; Score 40 ; DB 2 ; Length 91 ;

Best Local Similarity 66.7%; Pred. No. 0.55;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYGSKTLGRN 12
11:11111111
DB 41 FYGSKTLEYRN 52

RESULT 2

US-08-080-255-7
Sequence 7, Application US/08080255,
Patent No. 5487970
GENERAL INFORMATION:
APPLICANT: Rowley, Janet D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: DETECTING GENE REARRANGEMENTS AND
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/080,255
FILING DATE: 19930617
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:072/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 320-7200
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-080-255-7

Query Match 57.8%; Score 37; DB 1; Length 1400;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYGSKTLGRN 12
11:11111111
DB 61 FYGSKTLEYGRN 72

RESULT 3

US-08-465-713-7
Sequence 7, Application US/08465713
Patent No. 6121419
GENERAL INFORMATION:
APPLICANT: Rowley, Janet D.
APPLICANT: Diaz, Manuel O.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: DETECTING GENE REARRANGEMENTS AND
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433

CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,713
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/080,255
FILING DATE: 17 JUNE 1993
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:072/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 320-7200
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-465-713-7

Query Match 57.8%; Score 37; DB 4; Length 1400;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYGSKTLGRN 12
11:11111111
DB 61 FYGSKTLEYGRN 72

RESULT 4

PCT-US93-05857-7
Sequence 7, Application PC/TUS9305857
GENERAL INFORMATION:
APPLICANT: Board of Regents
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: GENE REARRANGEMENTS AND TRANSLOCATIONS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05857
FILING DATE: 19930617
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/900,689
FILING DATE: 17/06/92
ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:072/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 320-7200
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-05857-7

Query Match 57.8%; Score 37; DB 5; Length 1400;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FVMSKTLGRN 12
1 11:11111
DB 61 FTGSGTGRGN 72

RESULT 5
US-08-061-376-5
Sequence 5, Application US/08061376
Patent No. 6175000
GENERAL INFORMATION:
APPLICANT: Evans, Glen A.
APPLICANT: Diabali, Malek
APPLICANT: Selleri, Licia
APPLICANT: Parry, Pauline
TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 11023
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/061,376
FILING DATE: 13-MAY-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9387
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)546-4737
TELEFAX: (619)546-9392
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-061-376-5

Query Match 57.8%; Score 37; DB 4; Length 3969;
Best Local Similarity 66.7%; Pred. No. 16+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FVMSKTLGRN 12
1 11:11111
DB 838 FTGSGTGRGN 849

RESULT 6
US-08-403-866-5
Sequence 5, Application US/08403866
Patent No. 5643779
GENERAL INFORMATION:
APPLICANT: Ehrlich, Stanislaw
APPLICANT: Godon, Jean-Jacques
APPLICANT: Renault, Pierre
TITLE OF INVENTION: Nucleic acid coding for an alpha-acetolactate
NUMBER OF SEQUENCES: 16
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,866
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30, 727
REFERENCE/DOCKET NUMBER: 20747/30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1600
TELEFAX: (716) 263-1487
TELEX: 978450 (WUT)
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis subsp. lactis
INDIVIDUAL ISOLATE: ORF2
US-08-403-866-5

Query Match 56.2%; Score 36; DB 1; Length 259;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 MGSKTLGRN 12
11:11111
DB 129 MGAKSLGRN 138

RESULT 7
US-08-469-667-7
Sequence 7, Application US/08469667
Patent No. 5733748
GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gillfillan, Cecchi,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,667
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-667-7

Query Match 51.6%; Score 33; DB 1; Length 135;
Best Local Similarity 45.5%; Pred. No. 22;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 YMGSKTLGRN 12
:::|:|:|:
Db 92 FIGAKDLRGRS 102

RESULT 8
PCT-US95-07289-7
Sequence 7, Application PC/TUS9507289
GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bath, Gilfillan, Cecchi,
STREET: Stewart & Olstein
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07289
FILING DATE: 06-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-265
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-07289-7

Query Match 51.6%; Score 33; DB 5; Length 135;
Best Local Similarity 45.5%; Pred. No. 22;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 YMGSKTLGRN 12
:::|:|:|:
Db 92 FIGAKDLRGRS 102

RESULT 9
US-08-428-414A-2
Sequence 2, Application US/08428414A
Patent No. 5912166
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TITLE OF INVENTION: LEISHMANIASIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-APR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Radicek, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121.407
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 322 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-428-414A-2

Query Match 51.6%; Score 33; DB 2; Length 322;
Best Local Similarity 54.5%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMGSKTLGR 11
|:|:|:|:
Db 51 FYMGSKTLGR 61

RESULT 10
US-08-341-843B-4
Sequence 4, Application US/08341843B
Patent No. 5872225
GENERAL INFORMATION:
APPLICANT: Lemmon, Vance
TITLE OF INVENTION: A Method for Characterizing the
TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
Patent No. 5872225
TITLE OF INVENTION: the Nucleotide Sequence
TITLE OF INVENTION: Characterized Thereby

NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fay, Sharpe, Beall, Fagan,
ADDRESSEE: Minnich & McKee
STREET: 1100 Superior Avenue
STREET: Suite 700
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2518
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
MEDIUM TYPE: Storable
COMPUTER: Compaq Prolinea 5100e
OPERATING SYSTEM: DOS 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/341,843B
FILING DATE: No. 5872225ember 18, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/904,991
FILING DATE: June 26, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Minnich, Richard J.
REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: CWR 2 149-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861-5582
TELEFAX: (216) 980162
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 95
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acids
HYPOTHETICAL: Irrelevant
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: mouse
INDIVIDUAL ISOLATE: 8 day old mouse brain
IMMEDIATE SOURCE:
LIBRARY: lambda GT 10 and lambda GT11
CLONE: synthesis of several clones
PUBLICATION INFORMATION:
AUTHORS: Moos, M.
AUTHORS: Tacke, R.
AUTHORS: Scherer, H.
AUTHORS: Teplov, D.
AUTHORS: Fruh, K.
AUTHORS: Schachner, M.
TITLE: Neural adhesion molecule L1 is a
TITLE: member of the immunoglobulin
TITLE: superfamily with binding domains
TITLE: similar to fibronectin
JOURNAL: NATURE
VOLUME: 334
ISSUE:
PAGES: 701-703
DATE: 1988
US-08-341-843B-4

Query Match 50.0%; Score 32; DB 2; Length 95;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 YMSKTLGRN 12
DB 56 YSGSFTEGNN 66

RESULT 11
US-08-427-497E-9
Sequence 9, Application US/08427497E
Patent No. 5969124
GENERAL INFORMATION:
APPLICANT: Lemmon, Vance
TITLE OF INVENTION: A Method for Characterizing the
TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
Patent No. 5969124
TITLE OF INVENTION: the Nucleotide Sequence
TITLE OF INVENTION: Characterized Thereby
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fay, Sharpe, Beall, Fagan,
ADDRESSEE: Minnich & McKee
STREET: 1100 Superior Avenue
STREET: Suite 700
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2518
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
MEDIUM TYPE: Storable
COMPUTER: Compaq Prolinea 5100e
OPERATING SYSTEM: DOS 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,497E
FILING DATE: April 24, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/904,991
FILING DATE: June 26, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Minnich, Richard J.
REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861-5582
TELEFAX: (216) 241-1666
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 95
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acids
HYPOTHETICAL: Irrelevant
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: mouse
INDIVIDUAL ISOLATE: 8 day old mouse brain
IMMEDIATE SOURCE:
LIBRARY: lambda GT 10 and lambda GT11
CLONE: synthesis of several clones
PUBLICATION INFORMATION:
AUTHORS: Moos, M.
AUTHORS: Tacke, R.
AUTHORS: Scherer, H.
AUTHORS: Teplov, D.
AUTHORS: Fruh, K.
AUTHORS: Schachner, M.
TITLE: Neural adhesion molecule L1 is a
TITLE: member of the immunoglobulin
TITLE: superfamily with binding domains
TITLE: similar to fibronectin
JOURNAL: NATURE
VOLUME: 334
ISSUE:
PAGES: 701-703

DATE: 1988
US-08-427-497E-9

Query Match 50.0%; Score 32; DB 2; Length 95;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 YMGSKTLRGN 12
DB 56 YGSKFTIBGN 66

RESULT 12
US-09-405-035-1
; Sequence 1, Application US/09405035
; Patent No. 6184351
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Clifton, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/09/405,035
; CURRENT FILING DATE: 1999-09-24
; EARLIER APPLICATION NUMBER: 60/101,862
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-09-405-035-1

Query Match 50.0%; Score 32; DB 4; Length 140;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 YMGSKTLRG 10
DB 39 YGSKTKEG 47

RESULT 13
US-09-405-035-2
; Sequence 2, Application US/09405035
; Patent No. 6184351
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Clifton, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/09/405,035
; CURRENT FILING DATE: 1999-09-24
; EARLIER APPLICATION NUMBER: 60/101,862
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-09-405-035-2

Query Match 50.0%; Score 32; DB 4; Length 140;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 2 YMGSKTLRG 10

DB 39 YGSKTKEG 47

RESULT 14
US-09-405-035-3
; Sequence 3, Application US/09405035
; Patent No. 6184351
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Clifton, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/09/405,035
; CURRENT FILING DATE: 1999-09-24
; EARLIER APPLICATION NUMBER: 60/101,862
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-09-405-035-3

Query Match 50.0%; Score 32; DB 4; Length 140;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 YMGSKTLRG 10
DB 39 YGSKTKEG 47

RESULT 15
US-09-405-035-4
; Sequence 4, Application US/09405035
; Patent No. 6184351
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Clifton, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/09/405,035
; CURRENT FILING DATE: 1999-09-24
; EARLIER APPLICATION NUMBER: 60/101,862
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-09-405-035-4

Query Match 50.0%; Score 32; DB 4; Length 140;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 YMGSKTLRG 10
DB 39 YGSKTKEG 47

Search completed: June 4, 2001, 12:14:30
Job time: 223 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:41 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-24

Perfect score: 64

Sequence: 1 FYMGSKTLGRN 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	64	100.0	344	2	A41357	Fc gamma (Igg) rec
2	64	100.0	374	1	A39878	Fc gamma (Igg) rec
3	40	62.5	216	2	H64726	Yabp protein - Esc
4	40	62.5	404	2	I48471	Fc gamma (Igg) rec
5	40	62.5	404	2	A46480	Fc gamma (Igg) rec
6	39	60.9	379	1	E71162	hypothetical prote
7	38	59.4	276	2	H71147	hypothetical prote
8	37	57.8	212	2	F64940	hypothetical prote
9	37	57.8	1163	2	A56097	arylphorin-binding
10	37	57.8	3968	2	A44265	trichoxan homolog
11	36	56.2	259	2	F36889	leud 3'-region hyp
12	36	56.2	259	2	S35136	hypothetical prote
13	36	56.2	280	2	T18443	hypothetical prote
14	36	56.2	312	2	S49612	transposase - Esch
15	36	56.2	385	2	T07159	GTP-binding regula
16	36	56.2	389	2	A38302	pepsin (EC 3.4.23.
17	36	56.2	669	2	A47302	proline dehydrogen
18	36	56.2	1840	2	T30250	GTL protein - mous
19	35	54.7	599	2	D70104	DNA topoisomerase
20	35	54.7	726	2	D72392	cation transportin
21	35	54.7	847	2	A48228	beta-N-acetylhexos
22	34	53.1	266	2	S60674	hypothetical prote
23	34	53.1	286	2	T16681	hypothetical prote
24	34	53.1	292	2	B72659	deoxyhypusine synt
25	34	53.1	309	2	C83393	hypothetical prote
26	34	53.1	388	1	MMVZM1	major envelope ant
27	34	53.1	388	1	MMVZU2	major envelope ant
28	34	53.1	388	2	T30623	major envelope ant
29	34	53.1	527	2	T15606	hypothetical prote

30	34	53.1	531	2	S76041	hypothetical prote
31	34	53.1	736	2	S57961	dimethylamine dehy
32	34	53.1	771	2	H72410	hypothetical prote
33	33	51.6	107	2	T16521	hypothetical prote
34	33	51.6	165	2	E70486	hypothetical prote
35	33	51.6	167	2	T29565	hypothetical prote
36	33	51.6	236	2	T16164	hypothetical prote
37	33	51.6	249	2	S75500	short-chain alcoho
38	33	51.6	267	2	D72499	hypothetical prote
39	33	51.6	278	1	TPH0TW	tropoin T, slow s
40	33	51.6	291	2	S10867	early E4 34k prote
41	33	51.6	335	2	G64385	hypothetical prote
42	33	51.6	377	1	A44216	major envelope ant
43	33	51.6	379	2	A82806	integral membrane
44	33	51.6	383	2	A41545	pregnancy-specific
45	33	51.6	408	2	T18813	hypothetical prote

ALIGNMENTS

RESULT 1
Fc gamma (Igg) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000
C:Accession: A41357; S03019
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A:Title: Isolation and expression of functional high-affinity Fc receptor complementa
A:Reference number: A41357; M0ID:89100284
A:Accession: A41357
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: GB:X14355; GB:M21090; NID:931333; PIDN:CA432536.1; PID:931334
R:Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (A:Reference number: S03018; M0ID:89098339
A:Accession: S03019
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: EMBL:X14355; NID:931333; PIDN:CA432536.1; PID:931334
A:Note: the authors translated the codon ACT for residue 25 as Ala
C:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
F:117-170/Domain: immunoglobulin homology <IM>

Query Match 100.0%; Score 64; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 6; 9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12
DB 229 FYMGSKTLGRN 240

RESULT 2
A39878
Fc gamma (Igg) receptor I-A (high affinity) precursor - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: A39878; I70304; B41357; S03018; I57525
R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
J. Biol. Chem. 266, 13449-13455, 1991
A:Title: Gene organization of the human high affinity receptor for Igg, Fc gamma M1 (CD
A:Reference number: A39878; M0ID:91302383
A:Accession: A39878
A:Molecule type: DNA
A:Residues: 1-374 <VAN>

A:Cross-references: GB:M63830; GB:M63835; NID:g180277; PIDN:AAA5678.1; PID:g180279
 R.Porges, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: 155577; MUID:93053454
 A:Accession: 170304
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:U03418; NID:g184840; PIDN:AAA36049.1; PID:g292169
 R.Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: AA1357; MUID:89100284
 A:Accession: B41357
 A:Molecule type: mRNA
 A:Residues: 1-24,'S',26-337,'T',339-374 <ALL>
 A:Cross-references: GB:X14356; GB:M21091; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R.Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR)
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24,'S',26-337,'T',339-374 <AL2>
 A:Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R.Perez, C.; Metzgerh, J.; Benech, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A:Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
 15m.
 A:Reference number: 157525; MUID:93204964
 A:Accession: 157525
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RE2>
 A:Cross-references: GB:S57204; NID:g298692; PIDN:AA13887.1; PID:g4261587
 C:Genetics:
 A:Gene: GDB:FCGR1A; CD64
 A:Cross-references: GDB:135911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane F
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-292/Domain: extracellular #status predicted <EXT>
 F:117-170/Domain: immunoglobulin homology <IMM2>
 F:293-313/Domain: transmembrane #status predicted <TM>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent)#status predicte

Query Match 100.0%; Score 64; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 7.5e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYMSKTLRGRN 12
 DB 229 FYMSKTLRGRN 240

RESULT 3
 H64726
 Yabp protein - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 04-Mar-2000
 A:Accession: H64726
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: H64726
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-216 <BLAT>

A:Cross-references: GB:AE000116; GB:U00096; NID:g1786240; PIDN:AA073167.1; PID:g17862
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: yabp
 C:Superfamily: Escherichia coli yabp protein

Query Match 62.5%; Score 40; DB 2; Length 216;
 Best Local Similarity 54.5%; Pred. No. 2.4;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 YMSKTLRGRN 12
 DB 169 YGKKTMKCSN 179

RESULT 4
 I48471
 Fc gamma (IgG) receptor high affinity - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
 C:Accession: I48471
 R.Prins, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarth, P.M.; Wicker, L.S.; G
 Science 260, 695-698, 1993
 A:Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for
 A:Reference number: I48471; MUID:93242399
 A:Accession: I48471
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-336 <RES>
 A:Cross-references: EMBL:X70980; NID:g311748; PIDN:CAA50311.1; PID:g311749
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: immunoglobulin receptor
 F:128-180/Domain: immunoglobulin homology <IMM>

Query Match 62.5%; Score 40; DB 2; Length 336;
 Best Local Similarity 66.7%; Pred. No. 3.8;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMSKTLRGRN 12
 DB 243 FYGSKTLEYYRN 254

RESULT 5
 A46480
 Fc gamma (19C) receptor high affinity - mouse
 N:Alternate names: high affinity 19C receptor
 C:Species: Mus musculus (house mouse)
 C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
 C:Accession: A46480; A43511
 R.Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.
 J. Immunol. 148, 1570-1575, 1992
 A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI a
 A:Reference number: A46480; MUID:92166399
 A:Accession: A46480
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-404 <OSM>
 A:Note: sequence extracted from NCBI backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212,
 R:Seers, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.
 J. Immunol. 144, 371-378, 1990
 A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for
 A:Reference number: A43511; MUID:90111035
 A:Accession: A43511
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-404 <SEA>
 A:Cross-references: GB:M31314; NID:g200752; PIDN:AAA40056.1; PID:g200753
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: immunoglobulin receptor; transmembrane protein
 F:127-179/Domain: immunoglobulin homology <IMM>

Query Match 62.5%; Score 40; DB 2; Length 404;
Best Local Similarity 66.7%; Pred. No. 4.7;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 FYMGSKTLGRN 12
||:|||||
DB 238 FYVGSRTLEGRN 249

RESULT 6
E71162
hypothetical protein PH0498 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: E71162
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137
A:Accession: E71162
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <RAW>
A:Cross-references: GB:AP000002; NID:93236129; PIDN:BAA29586.1; PID:93256903
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Genbank
C:Genetics:
A:Gene: PH0498
C:Superfamily: hypothetical protein MJ0414

Query Match 60.9%; Score 39; DB 1; Length 379;
Best Local Similarity 58.3%; Pred. No. 6.9;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12
||:|||||
DB 273 FYIAEKRLGRN 284

RESULT 7
H71147
hypothetical protein PH0394 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C:Accession: H71147
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137
A:Accession: H71147
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-276 <RAW>
A:Cross-references: GB:AP000002; NID:93236129; PIDN:BAA29469.1; PID:dl030412; PID:932567
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Genbank
C:Genetics:
A:Gene: PH0394

Query Match 59.4%; Score 38; DB 2; Length 276;
Best Local Similarity 70.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 FYMGSKTLGRN 10
||:|||||
DB 264 FYAGSKVFRG 273

RESULT 8

F64940
hypothetical protein b1798 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 29-Sep-1999
C:Accession: F64940
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A:Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: F64940
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-212 <BLAT>
A:Cross-references: GB:AE000274; GB:U00096; NID:g1788089; PIDN:AAC74868.1; PID:g17880
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: hypothetical protein b1798

Query Match 57.8%; Score 37; DB 2; Length 212;
Best Local Similarity 53.3%; Pred. No. 9.4;
Matches 8; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 2 YMGSKTLGRN 12
||:|||||
DB 89 YLGSRTLYATLGRN 103

RESULT 9

A56097
arylphorin-binding protein precursor - flesh fly (Sarcophaga peregrina)
C:Species: Sarcophaga peregrina
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 17-Mar-1999
C:Accession: A56097
R:Chung, S.O.; Kubo, T.; Natori, S.
J. Biol. Chem. 270, 4624-4631, 1995
A:Title: Molecular cloning and sequencing of arylphorin-binding protein in protein gr
A:Reference number: A56097; MUID:95181458
A:Accession: A56097
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-163 <CHD>
A:Cross-references: GB:D29741

Query Match 57.8%; Score 37; DB 2; Length 1163;
Best Local Similarity 54.5%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YMGSKTLGRN 12
||:|||||
DB 464 YTGRTIMGRN 474

RESULT 10

A44265
titlthorax homolog HTX, version 2 - human
N:Alternate names: acute lymphoblastic leukemia gene 1 protein (ALL-1)
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C:Accession: A44265; A44264; I58112; I37165; I38485
R:Trachuk, D.C.; Kohler, S.; Cleary, M.L.
Cell 71, 691-700, 1992
A:Title: Involvement of a homolog of Drosophila titlthorax by 11q23 chromosomal transi
A:Reference number: A44265; MUID:93046667
A:Accession: A44265
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-3968 <TKA>
A:Cross-references: GB:L04284; NID:g184393; PIDN:AAA58669.1; PID:g184394
A:Note: sequence extracted from NCBI backbone (NCBIF:117729)

R.Gu. Y.; Nakamura, T.; Alder, H.; Prasad, R.; Cnaan, O.; Cimino, G.; Croce, C.M.; Can
Cell 71, 701-708, 1992
A:Title: The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 g
A:Reference number: A44264; MUID:9304668
A:Accession: A44264
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 63-316, 'GILLINSELEK', 327, 'Q', 329, 'VR', 332, 'DKEGTP', 340, 'T', 342, 'EDKTVVROSPR
546, 'LOISTSP', 2555-3554, 'N', 3556-3594, 'V', 3596-3899, 'A', 3901-3968 <GUT>
A:Cross-references: GB:L04731; NID:9339921
A>Note: Sequence extracted from NCBI backbone (NCBI:117779)
R.Djabbali, M.; Selleri, L.; Parry, P.; Bower, M.; Young, B.D.; Evans, G.A.
Nature Genet. 2, 113-118, 1992
A:Title: A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute
A:Reference number: 158112; MUID:93265134
A:Accession: 158112
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1317-1700, 'DD', 1703-1936, 'H', 1938-2180, 'S', 2182-2328 <DNA>
A:Cross-references: GB:L01986; NID:9307522; PIDN:AA92511.1; PID:953800
R:Marchalek, R.; Grell, J.; Lochner, K.; Nilsson, I.; Siegler, G.; Zweckbrunner, I.; Bed
Br. J. Haematol. 90, 308-320, 1995
A:Title: Molecular analysis of the chromosomal breakpoint and fusion transcripts in the
A:Reference number: 137165; MUID:95315013
A:Accession: 137165
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1212-1603, 'GTE', <MAR>
A:Cross-references: EMBL:X63604; NID:9897757; PIDN:CA58584.1; PID:9899268
A>Note: submitted to the EMBL/Genbank/DBJ databases by R. Marchalek, 20 December 1994
R.Gu. Y.; Alder, H.; Nakamura, T.; Schichman, S.A.; Prasad, R.; Cnaan, O.; Saito, H.;
Cancer Res. 54, 2327-2330, 1994
A:Title: Sequence analysis of the breakpoint cluster region in the ALL-1 gene involved in
A:Reference number: 138485
A:Accession: 138485
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1251-1486, 'G', 1488-1538 <RES>
A:Cross-references: EMBL:U04737; NID:9451554; PID:9451555
C:Genetics:
A:Gene: GDB:ML: HTX; ALL-1; HRX
A:Cross-references: GDB:128819; OMIM:159555
A:Map position: 11q23-11q23
A:Introns: 1338/1; 1362/3; 1406/3; 1444/3; 1493/3; 1525/3; 1566/1
A>Note: the list of introns is incomplete
C:Keywords: acute lymphoblastic leukemia; proto-oncogene; zinc finger
F:1434-1456/Region: zinc finger CCHC motif
F:1479-1506/Region: zinc finger CCHC motif
F:1527-1556/Region: zinc finger CCHC motif
F:1569-1596/Region: zinc finger CCHC motif
F:1873-1900/Region: zinc finger CCHC motif
F:1933-1955/Region: zinc finger CCHC motif

Query Match 57.8%; Score 37; DB 2; Length 3968;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
DB 839 FTGSGTGRGN 850

RESULT 11
F36889
leud 3 -region hypothetical protein - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 17-Mar-2000
C:Accession: F36889
R:Godon, J.J.; Delorme, C.; Bardowski, J.; Chopin, M.C.; Ehrlich, S.D.; Renault, P.
J. Bacteriol. 175, 4383-4390, 1993
A:Title: Gene inactivation in Lactococcus lactis: branched-chain amino acid biosynthesis
A:Reference number: A36889; MUID:93322316

A:Accession: F36889
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-259 <GOD>
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
C:Keywords: ATP; P-loop
F:19-219/Domain: ATP-binding motif A (P-loop)
F:36-43/Region: nucleotide-binding motif A (P-loop)

Query Match 56.2%; Score 36; DB 2; Length 259;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 MGSKTLGRN 12
DB 129 MGAKSLGRN 138

RESULT 12
S35136
hypothetical protein (leud 3' region) - Lactococcus lactis subsp. lactis
C:Species: Lactococcus lactis subsp. lactis
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C:Accession: S35136
R:Godon, J.J.; Chopin, M.C.; Ehrlich, S.D.
J. Bacteriol. 174, 6580-6589, 1992
A:Title: Branched-chain amino acid biosynthesis genes in Lactococcus lactis subsp. la
A:Reference number: S35132; MUID:93015710
A:Accession: S35136
A:Molecule type: DNA
A:Residues: 1-259 <GOD>
A:Cross-references: EMBL:M90761; NID:92565137; PIDN:AB81917.1; PID:92565155
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
C:Keywords: ATP; P-loop
F:19-219/Domain: ATP-binding motif A (P-loop)
F:36-43/Region: nucleotide-binding motif A (P-loop)

Query Match 56.2%; Score 36; DB 2; Length 259;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 MGSKTLGRN 12
DB 129 MGAKSLGRN 138

RESULT 13
T18443
hypothetical protein C0410w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18443
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18443
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-280 <LAW>
A:Cross-references: EMBL:Z98547; NID:e1325376; PID:e1407850; PIDN:CAB11120.2
C:Genetics:
A:Map position: 3
A:Introns: 19/1; 61/3; 87/2; 117/2; 145/3; 171/2
A>Note: C0410w

Query Match 56.2%; Score 36; DB 2; Length 280;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12

Db 98 FYYDNKVERGKN 109

RESULT 14

S49612
transposase - Escherichia coli
C:Species: Escherichia coli
C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 11-Jan-2000
C:Accession: S49612
R:Yavakoli, N.P.; Comanducci, A.; Dodd, H.; Bennett, P.M.
submitted to the EMBL Data Library, November 1994
A:Description: PUB2380 carries a putative transposable element (IS1294) capable of mediating transposition
A:Reference number: S49612
A:Accession: S49612
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <TAV>
A:Cross-references: EMBL:X82430; NID:g572684; PID:g572685
C:Genetics:
A:Gene: tnp1294
C:Superfamily: Escherichia IS91 transposase: tnpA

Query Match 56.2% Score 36; DB 2; Length 312;
Best Local Similarity 63.6%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 YMGSKTLGRN 12
Db 198 YMGSKTAGGRN 208

RESULT 15

T07159
GTP-binding regulatory protein alpha chain 2 - soybean
C:Species: Glycine max (soybean)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
C:Accession: T07159
R:Kim, W.Y.; Cheong, N.E.; Lee, D.C.; Je, D.Y.; Bahk, J.D.; Cho, M.J.; Lee, S.Y.
Plant Physiol. 108, 1315-1316, 1995
A:Title: Cloning and sequencing analysis of a full-length cDNA encoding a G protein alpha subunit
A:Reference number: Z15966; MUID:95357428
A:Accession: T07159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-385 <KIM>
A:Cross-references: EMBL:L27418; NID:g439616; PID:AAA9517.1; PID:g439617
A:Experimental source: cultivar Williams
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; signal transduction

Query Match 56.2% Score 36; DB 2; Length 385;
Best Local Similarity 58.3%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
Db 367 FKIGDETLRRRN 378

Search completed: June 4, 2001, 12:15:42
Job time: 275 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:04 ; Search time 37.56 Seconds

(without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-24

Sequence: 1 FVMSKTLKRN 12

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	64	100.0	FCG1_HUMAN	P12314 homo sapien
2	40	62.5	YABP_ECOLI	P39220 escherichia
3	40	62.5	FCG1_MOUSE	P26151 mus musculu
4	39	60.9	RIAA0_LEITN	P39097 leishmania
5	38	59.4	YPC4_CAEEL	Q1181 caenorhabdi
6	37	57.8	YEAS_ECOLI	P76249 escherichia
7	37	57.8	E434_ADE09	P89083 human adeno
8	37	57.8	MTM1_MOUSE	09225 mus musculu
9	37	57.8	HRX_HUMAN	00164 homo sapien
10	36	56.2	YIUD_LACLA	00251 lactococcus
11	36	56.2	GBA1_SOYBN	P49084 glycine max
12	36	56.2	PROD_DROME	Q04499 drosophila
13	36	56.2	TREB_NEUCR	042783 neuropept
14	35	54.7	PARB_BORBU	059189 borrelia bu
15	35	54.7	HEX_VIBVU	004786 vibrio vuln
16	34	53.1	DHYS_AERPE	09272 aeropyrum p
17	34	53.1	VENW_MCV1	P26572 molluscum p
18	34	53.1	VENW_MCV2	P25392 molluscum c
19	34	53.1	NDOU_RHOC	P50339 rhodobacter
20	34	53.1	DHDM_HYPSX	048303 hyphomicrob
21	33	51.6	TDR1_HUMAN	P13805 homo sapien
22	33	51.6	E434_ADE12	P36710 human adeno
23	33	51.6	RIAA0_LEICH	P39096 leishmania
24	33	51.6	Y687_METUA	058100 methanococ
25	33	51.6	VENW_POWPY	P36316 fowlpox vir
26	33	51.6	PAG1_SHEEP	028315 ovis aries
27	33	51.6	EFIG_ARTSA	053728 artemia sal
28	33	51.6	PLD_STRAT	009698 schizontomye
29	33	51.6	KALB_SCHPO	009698 schizontomye
30	33	51.6	DIG2_RAT	063622 rattus norv
31	33	51.6	RESA_PLAFA	P13830 plasmodium
32	33	51.6	RPOD_SPIOL	P11704 spinacia ol
33	32	50.0	URE2_BACSU	P71035 bacillus su

34	32	50.0	134	1	SYUB_BOVIN	P33567 bos taurus
35	32	50.0	137	1	SYUB_HUMAN	O16143 homo sapien
36	32	50.0	137	1	SYUB_RAT	O63754 rattus norv
37	32	50.0	140	1	SYU1_RAT	P37377 rattus norv
38	32	50.0	140	1	SYUA_HUMAN	P37840 homo sapien
39	32	50.0	140	1	SYUA_MOUSE	O55042 mus musculu
40	32	50.0	149	1	SYU2_RAT	P37378 rattus norv
41	32	50.0	242	1	PMU2_MOUSE	O922m7 mus musculu
42	32	50.0	246	1	PMU2_HUMAN	O15305 homo sapien
43	32	50.0	257	1	E434_ADE40	O6485 human adeno
44	32	50.0	431	1	UROK_HUMAN	P00749 homo sapien
45	32	50.0	505	1	SYE_CHLPS	O06560 chlamydia p

ALIGNMENTS

RESULT 1
ID FCG1_HUMAN STANDARD: PRT: 374 AA.
AC P12314; P12315;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I) (CD64).
GN FCGRIA OR FCGRI OR FCGI OR IGFRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;
RX MEDLINE=89100284; PubMed=2974947;
RA Allen J.M., Seed B.;
RT "Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcRI)."
RL Nucleic Acids Res. 16:11824-11824(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89100284; PubMed=2974947;
RA Allen J.M., Seed B.;
RT "Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcRI)."
RL Nucleic Acids Res. 16:11824-11824(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89100284; PubMed=2974947;
RA Allen J.M., Seed B.;
RT "Isolation and expression of functional high-affinity Fc receptor complementary DNAs."
RL Science 243:378-381(1989).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH AFFINITY RECEPTOR.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide C64 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/c64.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC
CC EMBL: X14356; CAA32537.1;
CC EMBL: X14355; CAA32536.1;
CC PIR: S03018; S03018.
CC PIR: S03019; S03019.
CC PIR: A41357; A41357.
CC PIR: B41357; B41357.
CC HSSP: P12319; IALT.
CC MIM: 146760;
CC Interpro: IPR003006;
CC Pfam: PR00047; 1g; 3.

KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KM Immunoglobulin domain; Alternative splicing; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 374
 FT DOMAIN 22 292
 FT TRANSMEM 293 313
 FT DOMAIN 314 374
 FT DOMAIN 36 92
 FT DOMAIN 117 175
 FT DOMAIN 205 267
 FT DISULFID 43 85
 FT DISULFID 124 168
 FT DISULFID 212 260
 FT CARBOHYD 59 59
 FT CARBOHYD 78 78
 FT CARBOHYD 152 152
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT VARSPLIC 333 374
 FT VARIANT 26 26
 FT VARIANT 59 59
 FT CONFLICT 25 25
 FT SEQUENCE 374 AA; 42605 MW; 2C2A8103EFC16E6 CRC64;
 S -> T (IN CAA32536)
 /FTID=VAR.003954.
 N->I.V.
 /FTID=VAR.003953.
 L->T.
 T->GQALEAPTOGCA (IN ISOFORM B).
 HEKVTSSLDQDRHEELKCOEKEDELDQGVHKKPEOGA
 Query Match 100.0%; Score 64; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 2.8e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FYMGSKTLGRN 12
 DB 229 FYMGSKTLGRN 240
 ID ID
 AC P39220; STANDARD; PRT; 216 AA.
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHEICAL 24.8 KDA PROTEIN IN DJIA-KLDA INTERGENIC REGION.
 GN YARP.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=562;
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RA MEDLINE=92334977; Pubmed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RA "Systematic sequencing of the Escherichia coli genome: analysis of
 the 0.2-4 min region.";
 RA Nucleic Acids Res. 20:3305-3308(1992).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; Pubmed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RA Science 277:1233-1238(1997).
 RL [3]

RP IDENTIFICATION.
 RA Rudd K.E.;
 RL Unpublished observations (NOV-1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D10483; NOT ANNOTATED_CDS.
 DR EMBL: AE000116; AAC73167.1; -
 DR Ecogene; EGI2610; yabp.
 DR Hypothetical protein.
 KW SEQUENCE 216 AA; 24791 MW; A15D252B84B50621 CRC64;
 Query Match 62.5%; Score 40; DB 1; Length 216;
 Best Local Similarity 54.5%; Pred. No. 1;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 2 FYMGSKTLGRN 12
 DB 169 YLKKTKMGSN 179
 ID ID
 AC P26151; STANDARD; PRT; 404 AA.
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA
 DE RI) (FCRI) (IGG FC RECEPTOR I).
 GN FCGR1 OR FCGR1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90111035; Pubmed=2136886;
 RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RA "Molecular cloning and expression of the mouse high affinity Fc
 RA receptor for IgG.";
 RA J. Immunol. 144:371-378(1990).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=92163399; Pubmed=1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RA "Structure and mapping of the gene encoding mouse high affinity Fc
 RA gamma RI and chromosomal location of the human Fc gamma RI gene.";
 RA J. Immunol. 148:1570-1575(1992).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH
 CC AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M31314; AAA40056.1; -
 DR PIR: A43511; A43511.

DR	PIR:	A46480;	A46480.
DR	HSSP:	P12319;	IALT.
DR	MGD:	MGI:95498;	Fcgr1.
DR	InterPro:	IPRO03006;	-
DR	Pfam:	PF00047;	Ig; 3.
KW	IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;		
KM	Immunoglobulin domain.		
FT	SIGNAL	1	24
FT	CHAIN	25	404
FT	DOMAIN	25	297
FT	TRANSMEM	298	320
FT	DOMAIN	321	404
FT	DOMAIN	46	66
FT	DOMAIN	67	124
FT	DOMAIN	154	216
FT	CARBOHYD	28	28
FT	CARBOHYD	48	48
FT	CARBOHYD	69	69
FT	CARBOHYD	168	168
FT	CARBOHYD	249	249
SO	SEQUENCE	404 AA;	44887 MW; IC4F0033842767E7 CRC64;
Query Match			
	Best Local Similarity	62.5%;	Score 40; DB 1; Length 404;
	Matches	8; Conservative	1; Pred. NO. 2;
			Gaps 0;
QY	1 FYMGSKTLGRGN 12		
DB	238 FVGSKILRYRN 249		
RESULT 4			
ID	R1AO LEIIN	STANDARD;	PRT; 323 AA.
AC	P39097;		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-FEB-1995 (Rel. 31, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	60S ACIDIC RIBOSOMAL PROTEIN P0.		
LN	LPO-A AND LPO-B.		
OS	Leishmania infantum.		
OC	Eukaryote; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.		
OX	NCBI_TaxID=5671;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRATN-LEM 75;		
RX	MEDLINE=94088674; PubMed=8264730;		
RA	Soto M., Requena J.M., Alonso C.;		
RT	"Isolation, characterization and analysis of the expression of the		
RT	Leishmania ribosomal P0 protein genes."		
RL	Mol. Biochem. Parasitol. 61:265-274(1993).		
CC	-1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT		
CC	OF E-COLI PROTEIN L10.		
CC	-1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH		
CC	DIMERS OF P1 AND P2.		
CC	-1- SIMILARITY: BELONGS TO THE LIOP FAMILY OF RIBOSOMAL PROTEINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its use by		
CC	non-profit institutions as long as its content is in no way used for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to licenses@isb-sib.ch).		
CC	-----		
DR	EMBL: X72714; CAAS1264.1; -		
DR	EMBL: X72714; CAAS1263.1; -		
DR	InterPro: IPR001790; -		
DR	InterPro: IPR001813; -		
DR	Pfam: PF00428; 60S-Ribosomal; 1.		
DR	Pfam: PF00466; Ribosomal_L10; 1.		

```

KW Ribosomal protein; Phosphorylation.
SO SEQUENCE 323 AA; 34771 MW; 14967BD7A439D69E CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 1; Length 323;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FYMSGKTLGR 11
   |||||:
Db 51 FFMGKMTLGRK 61

RESULT 5
YPCA_CAEEL STANDARD; PRT; 371 AA.
AC 011181;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHEETICAL 41.9. KDA PROTEIN C05D10.4 IN CHROMOSOME III.
GN C05D10.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Du Z.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U13645; AAA20988.1;
DR Wormpep; C05D10.4; CE01127.
DR Hypothetical protein.
SQ SEQUENCE 371 AA; 41906 MW; 6CE79B87D8DD390C CRC64;

Query Match
Best Local Similarity 59.4%; Score 38; DB 1; Length 371;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMSGKTLGRN 12
   |||||:
Db 25 FFMGKMTLGRN-36

RESULT 6
YEAS_ECOLI STANDARD; PRT; 212 AA.
AC P76249; C007971; C00769;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHEETICAL 23.2 KDA PROTEIN IN GAPA-RND INTERGENIC REGION.
GN YEAS.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
DE MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

```

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-97251358; PubMed-9097040;
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
 RA Mizouchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
 RA Nashimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y.,
 RA Sivasubraman S., Tagami H., Takeda J., Takemoto K., Wada C.,
 RA Yamamoto Y., Horuchi T.;
 RT "A 460-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 40.1-50.0 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE RHT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF000274; AAC74868.1;
 DR EMBL: D90823; BAA1593.1;
 DR EMBL: D90824; BAA15602.1;
 DR Ecogene; EGI3505; Yeas.
 DR InterPro; IPR001123;
 DR Pfam; PF01810; Lyse; 1.
 DR Hypothetical protein; Transmembrane.
 KW TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 49 69 POTENTIAL.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 153 173 POTENTIAL.
 FT TRANSMEM 188 208 POTENTIAL.
 SO SEQUENCE 212 AA; 23200 MW; 342E0DF348C9AD9A CRC64;

Query Match 57.8%; Score 37; DB 1; Length 212;
 Best Local Similarity 53.3%; Pred. No. 3.9;
 Matches 8; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

OY 2 YMGSK---TLGRN 12
 Db 89 YIGSKILVATLGRN 103

RESULT 7
 ID E434-ADE09 STANDARD; PRT; 292 AA.
 AC P89083;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 DE EARLY E4 34 KDA PROTEIN.
 OS Human adenovirus type 9.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OC NCBI_TaxID=10527;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96327849; PubMed=8738606;
 RA Javier R., Shenk T.;
 RT "Mammary tumors induced by human adenovirus type 9: a role for the
 RT viral early region 4 gene.";
 RL Breast Cancer Res. Treat. 39:57-67(1996).
 CC -1- SIMILARITY: BELONGS TO THE ADENOVIRUS E4 30 TO 34 KDA PROTEIN

CC FAMILY
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S82508; AAB37508.2;
 DR Early protein.
 KW SEQUENCE 292 AA; 34082 MW; 1F0B15EDFE2A2A1E CRC64;

Query Match 57.8%; Score 37; DB 1; Length 292;
 Best Local Similarity 63.6%; Pred. No. 5.5;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 YMGSKTLGRN 12
 Db 165 YMGSVFMGRH 175

RESULT 8
 ID MTM1_MOUSE STANDARD; PRT; 603 AA.
 AC Q92ZC5;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MYOTUBULARIN (EC 3.1.3.48).
 GN MTM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98409499; PubMed=9736772;
 RA Laporte J., Blondeau F., Bui-Bello A., Tentler D., Kretz C., Dahl N.,
 RA Mandel J.L.;
 RT "Characterization of the myotubularin dual specificity phosphatase
 RT gene family from yeast to human.";
 RL Hum. Mol. Genet. 7:1703-1712(1998).
 CC -1- FUNCTION: DUAL-SPECIFICITY PHOSPHATASE THAT ACTS ON BOTH
 CC PHOSPHOTYROSINE AND PHOSPHOSERINE. COULD BE INVOLVED IN A SIGNAL
 CC TRANSDUCTION PATHWAY NECESSARY FOR LATE MYOGENESIS, ALTHOUGH ITS
 CC SUBSTRATES EXPRESSION SUGGESTS A WIDER FUNCTION.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. MYOTUBULARIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF073996; AAC77821.1;
 DR MGD; MGI:1099452; Mtm1.
 DR InterPro; IPR000387;
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1;
 DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1;
 KW Hydrolase.
 FT ACT_SITE 375 375 BY SIMILARITY.
 SO SEQUENCE 603 AA; 69594 MW; 65929312F73FE2C7 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 603;
 Best Local Similarity 63.6%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 YMGSKTLRGN 12
DB 95 YMGCATSRGEN 105

RESULT 9
HRX_HUMAN

ID HRX_HUMAN STANDARD: PRT: 3969 AA.
AC 003164; Q14845; Q16364; Q13743; Q13744; Q9DMA3;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ZINC FINGER PROTEIN HRX (ALL-1) (TRITHORAX-LIKE PROTEIN).
GN HMO OR HRX OR ALL1 OR TRX1 OR HTRX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

[1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93046667; PubMed-1423624;
RT Kachuk D.C., Kohler S., Cleary M.L.;
RT "Involvement of a homolog of Drosophila trithorax by 11q23
RT chromosomal translocations in acute leukemias.";
RL Cell 71:691-700(1992).

[2]
RP SEQUENCE FROM N.A.
RX MEDLINE-96290553; PubMed-8703835;
RT Nilsson I., Loechner K., Sieglar G., Grell J., Beck J.D., Fey G.H.,
RT Marschalek R.;
RT "Exon/intron structure of the human ALL-1 (MLL) gene involved in
RT translocations to chromosomal region 11q23 and acute leukemias.";
RL Br. J. Haematol. 93:965-972(1996).

[3]
RP SEQUENCE OF 1-1909 FROM N.A.
RX MEDLINE-93390935; PubMed-8378076;
RT Yamamoto K., Seto M., Komatsu H., Iida S., Akao Y., Kojima S.,
RT Kodera Y., Nakazawa S., Ariyoshi Y., Takahashi T., Ueda K.;
RT "Two distinct portions of Lfng19/ENL at 19p13 are involved in t(11;19)
RT leukemia.";
RL Oncogene 8:2617-2625(1993).

[4]
RP SEQUENCE OF 1317-2328 FROM N.A.
RC TISSUE-Brain;
RX DJabali M., Selleri L., Parry P., Bower M., Young B.D., Evans G.A.;
RT "A trithorax-like gene is interrupted by chromosome 11q23
RT translocations in acute leukemias.";
RL Nat. Genet. 2:113-118(1992).

[5]
RP SEQUENCE OF 1251-1538 FROM N.A.
RX MEDLINE-94215165; PubMed-8162575;
RT Gu Y., Alder H., Nakamura T., Schlichman S.A., Prasad R., Canaan O.,
RT Saito H., Croce C.M., Canaan E.;
RT "Sequence analysis of the breakpoint cluster region in the ALL-1 gene
RT involved in acute leukemia.";
RL Cancer Res. 54:2326-2330(1994).

[6]
RP SEQUENCE OF 1251-1654 FROM N.A. (ISOFORM 14P-18B).
RX MEDLINE-95322025; PubMed-7598802;
RT Mbangkolo D., Burnett R., McCabe N., Thirman M., Gill H., Yu H.,
RT Rowley J.D., Diaz M.O.;
RT "The human MLL gene: nucleotide sequence, homology to the Drosophila
RT trit-zinc-finger domain, and alternative splicing.";
RL DNA Cell Biol. 14:475-483(1995).

[7]
RP SEQUENCE OF 1212-1603 FROM N.A.
RX MEDLINE-95315013; PubMed-7794749;
RT Marschalek R., Grell J., Loechner K., Nilsson I., Sieglar G.,
RT Zwechbrunner I., Beck J.D., Fey G.H.;
RT "Molecular analysis of the chromosomal breakpoint and fusion
RT transcripts in the acute lymphoblastic SEM cell line with chromosomal

RT translocation t(4;11).";
RL Br. J. Haematol. 90:308-320(1995).

[8]
RP SEQUENCE OF 1421-1540 FROM N.A.
RX MEDLINE-94020842; PubMed-8414518;
RT Forster A., Rabbits T.H.;
RT "A method for identifying genes within yeast artificial chromosomes:
RT application to isolation of MLL fusion cDNAs from acute leukemia
RT translocations.";
RL Oncogene 8:3157-3160(1993).

[9]
RP CHROMOSOMAL TRANSLOCATION WITH GAST.
RX MEDLINE-20183971; PubMed-10706619;
RT Megonigal M.D., Cheung N.-K.V., Kappaport E.F., Nowell P.C.,
RA Williams T.M., Lange B.J., Felix C.A.;
RT "Detection of leukemia-associated MLL-GAST translocation early during
RT chemotherapy with DNA topoisomerase II inhibitors.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2814-2819(2000).

CC -1- FUNCTION: POSSIBLY ACTS AS A TRANSCRIPTIONAL REGULATORY FACTOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: HEART, LUNG, BRAIN AND T AND B LYMPHOCYTES.
CC -1- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY CHROMOSOMAL TRANSLOCATIONS
CC T(11;19)(Q23;P13.3) THAT INVOLVES MLL AND MLLT7/AFX1;
CC T(4;11)(Q21;Q23) THAT INVOLVES MLL AND MLLT2/AF4; T(9;11)(P22;Q23)
CC THAT INVOLVES MLL AND MLLT3/AF9; T(6;11)(Q27;Q23) THAT INVOLVES
CC MLL AND MLLT4/AF6; T(11;17)(Q23;Q21) THAT INVOLVES MLL AND
CC MLLT6/AF17; T(X;11)(Q13;Q23) THAT INVOLVES MLL AND MLLT10/AF10;
CC T(10;11)(P12;Q23) THAT INVOLVES MLL AND MLLT10/AF10;
CC T(1;11)(Q21;Q23) THAT INVOLVES MLL AND AF10; T(11;19)(Q23;P13.3)
CC THAT INVOLVES MLL AND ELL; AND T(11;19)(Q23;P23) THAT INVOLVES MLL
CC AND GAST.
CC -1- SIMILARITY: BELONGS TO THE TRITHORAX FAMILY OF TRANSCRIPTION
CC FACTORS.
CC -1- SIMILARITY: CONTAINS 1 BROMODOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 PHD ZINC-FINGER DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: L04284; AAA58669.1; -
CC EMBL: 269744; CAA93625.1; -
CC EMBL: 269745; CAA93625.1; JOINED.
CC EMBL: 269746; CAA93625.1; JOINED.
CC EMBL: 269747; CAA93625.1; JOINED.
CC EMBL: 269748; CAA93625.1; JOINED.
CC EMBL: 269749; CAA93625.1; JOINED.
CC EMBL: 269750; CAA93625.1; JOINED.
CC EMBL: 269751; CAA93625.1; JOINED.
CC EMBL: 269752; CAA93625.1; JOINED.
CC EMBL: 269753; CAA93625.1; JOINED.
CC EMBL: 269754; CAA93625.1; JOINED.
CC EMBL: 269755; CAA93625.1; JOINED.
CC EMBL: 269756; CAA93625.1; JOINED.
CC EMBL: 269757; CAA93625.1; JOINED.
CC EMBL: 269758; CAA93625.1; JOINED.
CC EMBL: 269759; CAA93625.1; JOINED.
CC EMBL: 269760; CAA93625.1; JOINED.
CC EMBL: 269761; CAA93625.1; JOINED.
CC EMBL: 269762; CAA93625.1; JOINED.
CC EMBL: 269763; CAA93625.1; JOINED.
CC EMBL: 269764; CAA93625.1; JOINED.
CC EMBL: 269765; CAA93625.1; JOINED.
CC EMBL: 269766; CAA93625.1; JOINED.
CC EMBL: 269767; CAA93625.1; JOINED.
CC EMBL: 269768; CAA93625.1; JOINED.
CC EMBL: 269769; CAA93625.1; JOINED.

RESULT 10

ID	YLUD	IACLA	STANDARD	PRT	259 AA
AC	002151				
DT	01-JUL-1993	(Rel. 26, Created)			
DT	01-JUL-1993	(Rel. 26, Last sequence update)			
DT	15-DEC-1998	(Rel. 37, Last annotation update)			
DE	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN IN LEUD 3' REGION.				
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).				
OC	Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				
CC	Lactococcus				
CC	NCBI_TaxID=1360;				
CC	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-NCDO 2118;				
RX	MEDLINE=93015710; PubMed=1400210;				
RT	Gordon J. Y., Chopin M. -C., Ehrlich S.D.;				
RT	"Branched-chain amino acid biosynthesis genes in Lactococcus lactis				
RL	subsp. lactis."				
RL	J. Bacteriol. 174:6580-6589(1992).				
CC	-1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY				
CC	(ABC TRANSPORTERS).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as it is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL: U92974; AAB81917.1;				
DR	PIR: S35136; S35136.				
DR	InterPro: IPR001617;				
DR	Pfam: PF00005; ABC_tran; 1.				
DR	PROSITE: PS00211; ABC_TRANSPORTER; 1.				
KW	Hypothetical protein; ATP-binding; Transport.				
FT	NP_BIND 36				
FT	ATP (BY SIMILARITY).				
SO	SEQUENCE 259 AA; 29713 MW; 49ED150F6A15169 CRC64;				
QY	3 MGSTINGRN 12.				
Db	129 MGAKSLIGN 138				
RESULT 11					
GBAL	SOYBN				
ID	GBAL_SOYBN	STANDARD;	PRT;	385 AA.	
AC	P49084;				
DT	01-FEB-1996 (Rel. 33, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	GRANIE NUCLEOTIDE-BINDING PROTEIN ALPHA-1 SUBUNIT (GP-ALPHA-1).				
GN	GPAL OR GAL.				
OS	Glycine max (Soybean).				
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;				
OC	Magnoliophyta; eudicotyledons: core eudicots: Rosidae; eustosids I;				
OC	Fabales; Fabaceae; Papilionoideae; Glycine.				
OX	NCBI_TaxID=3847;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-CV. WILLIAMS;				
RX	MEDLINE=95357428; PubMed=7630957;				
RA	Kim W.Y., Cheong N.E., Lee D.C., Je D.Y., Baik J.D., Cho M.J.,				
RA	Lee S.Y.;				
RT	"Cloning and sequencing analysis of a full-length cDNA encoding a G				
RT	protein alpha subunit SGAL from soybean."				
RL	Plant Physiol. 108:1315-1316(1995).				
CC	-1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE				

CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -1- SUBUNIT: 6 PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY, SUBFAMILY 2 (G(S)).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L27418; AAA99517.1;
 CC DR HSSP: P10824; IAS3.
 CC DR InterPro: IPR001019;
 CC DR InterPro: IPR002976;
 CC DR Pfam: PF00503; G-alpha: 1.
 CC DR PRINTS: PR00318; GPROTEINA.
 CC DR PRINTS: PR01242; GPROTEINAPL.
 CC DR GTP-binding; Transducer; ADP-ribosylation.
 CC KW NP-BIND 46 53 GTP (BY SIMILARITY).
 CC FT NP-BIND 219 223 GTP (BY SIMILARITY).
 CC FT NP-BIND 288 291 GTP (BY SIMILARITY).
 CC FT MOD_RES 191 191 ADP-RIBOSYL(1) (BY ACTION OF CTX).
 CC SQ SEQUENCE 385 AA; 44919 MW; 65547803630FFAA CRC64;

Query Match 56.2%; Score 36; DB 1; Length 385;
 Best Local Similarity 58.3%; Pred. No. 12;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
 DB 367 FKLGETLRRN 378
 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 12
 PROD_DROME STANDARD; PRT; 669 AA.
 AC 004459; 061349;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROLINE OXIDASE, MITOCHONDRIAL PRECURSOR (EC 1.5.3.-) (PROLINE
 DE DEHYDROGENASE) (PROTEIN SLUGGISH-A).
 GN SLGA OR SLG.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CANTON-S;
 RC MEDLINE=93219408; PubMed=8096642;
 RA Hayward D.C., Delaney S.J., Campbell H.D., Ghysen A., Benzer S.,
 RA Kasprzak A.B., Cotsell J.N., Young I.G., Miklos G.L.G.;
 RT "The sluggish-A gene of Drosophila melanogaster is expressed in the
 RT nervous system and encodes proline oxidase, a mitochondrial enzyme
 RT involved in glutamate biosynthesis";
 RT Proc. Natl. Acad. Sci. U.S.A. 90:2979-2983(1993).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CANTON-S;
 RC MEDLINE=98188272; PubMed=9520435;
 RA Maleszka R., de Couet H.G., Miklos G.L.G.;
 RT "Data transferability from model organisms to human beings: insights
 RT from functional genomics of the flightless region of Drosophila";
 RT Proc. Natl. Acad. Sci. U.S.A. 95:3731-3736(1998).
 CC -1- FUNCTION: CONVERTS PROLINE TO DELTA-1-PYRROLINE-5-CARBOXYLATE.
 CC REDUCED PROLINE OXIDASE ACTIVITY PRODUCES SLUGGISH BEHAVIOR.

CC -1- PATHWAY: FIRST STEP IN THE CONVERSION FROM PROLINE TO GLUTAMATE.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN DEVELOPING NERVOUS SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN DEVELOPING EMBRYO AS WELL AS IN
 CC ADULT.
 CC -1- SIMILARITY: BELONGS TO THE PROLINE OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L07330; AAA02748.1;
 CC DR EMBL: AF01777; AAC28410.1;
 CC DR PIR: A47302; A47302.
 CC DR Flybase: FBgn003423; slga.
 CC DR InterPro: IPR002872;
 CC DR Pfam: PF01619; Pro_dh; 1.
 CC DR Oxidoreductase; Proline metabolism; Mitochondrion; Transl. peptide.
 CC KW TRANSIT 1 30 MITOCHONDRION (POTENTIAL).
 CC FT CHAIN 31 669 PROLINE OXIDASE.
 CC FT CONFLICT 42 42 N -> S (IN REF. 2).
 CC FT CONFLICT 147 149 NLY -> KLF (IN REF. 2).
 CC SQ SEQUENCE 669 AA; 75965 MW; D9C37A99DA8EC705 CRC64;

Query Match 56.2%; Score 36; DB 1; Length 669;
 Best Local Similarity 60.0%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 10
 DB 491 FYMGKTLGRN 500
 1 1 1 1 1 1 1 1 1 1

RESULT 13
 TREB_NEUCR STANDARD; PRT; 728 AA.
 AC 042783;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE NEUTRAL TREHALASE (EC 3.2.1.28) (ALPHA, ALPHA-TREHALASE)
 DE (ALPHA, ALPHA-TREHALOSE GLUCOHYDROLASE).
 GN TREB.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A;
 RA D'Entfer C., Bonini B., Zapella P.D.A., Fontaine T., da Silva A.M.,
 RA Terenzi H.F.;
 RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POSSIBLY INVOLVED IN MOBILIZATION OF CYTOSOLIC
 CC TREHALOSE.
 CC -1- CATALYTIC ACTIVITY: ALPHA, ALPHA-TREHALOSE + H(2)O -> 2 D-GLUCOSE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 37 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF044218; AAC01744.1;

DR InterPro: IPR001661; -
DR Pfam: PF01204; Trehalase 1.
DR PRINTS: PR00744; GLYHYDRASE37.
DR PROSITE: PS00927; TREHALASE_1; 1.
DR PROSITE: PS00928; TREHALASE_2; 1.
DR Hydrolase: Glycosidase.
SQ SEQUENCE 728 AA; 84121 MW; 04DD74D1C2E6154D CRC64;

Query Match 56.2%; Score 36; DB 1; Length 728;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 GSKTLGR 11
| | | | |
DB 271 GSKTLGR 278

RESULT 14

PARE_BORBU STANDARD; PRT; 599 AA.
AC Q59189;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TOPOISOMERASE IV SUBUNIT B (EC.5.99.1.-).
GN PARE OR BB0036.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE=9806543; Pubmed=9403685;

RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb O.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Matthey L., McDonald L., Artach P., Bowman C.,
RA Gierlicki S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RA "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi.";
RL Nature 390:580-586(1997).
RN [2]
RP SEQUENCE OF 1-83 FROM N.A.
RC STRAIN=212;
RX MEDLINE=9511614; Pubmed=7812434;
RA Ojalimi C., Davidson B.E., Saint-Girons I., Old I.G.;
RA "Conservation of gene arrangement and an unusual organization of rRNA
genes in the linear chromosomes of the Lyme disease spirochetes
Borrelia burgdorferi, B. garinii and B. afzelii.";
RL Microbiology 140:2931-2940(1994).

CC -1- FUNCTION: TOPOISOMERASE IV IS ESSENTIAL FOR CHROMOSOME
CC SEGREGATION. IT HAS RELAXATION OF SUPERCOILED DNA ACTIVITY.
CC PERFORMS THE DECATENATION EVENTS REQUIRED DURING THE REPLICATION
CC OF A CIRCULAR DNA MOLECULE (BY SIMILARITY).
CC -1- SUBUNIT: COMPOSED OF TWO SUBUNITS: PARC AND PARE.
CC -1- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: AE001117; AAC66418.1; -
CC EMBL: U32861; AAC41408.1; -
CC TIGR: BB0036; -
DR InterPro: IPR001241; -

DR Pfam: PF00204; DNA topoisom II; 1.
DR PRINTS: PR00418; TP12FAMILY.
DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
DR Isomerase: Topoisomerase; ATP-binding.
SQ SEQUENCE 599 AA; 68774 MW; B5901F17B1CC7721 CRC64;

Query Match 54.7%; Score 35; DB 1; Length 599;
Best Local Similarity 58.3%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
| | | | |
DB 581 FYMGNTPTERRN 592

RESULT 15

HEX_VIBVU STANDARD; PRT; 847 AA.
AC Q04786;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BETA-HEXOSAMINIDASE (EC 3.2.1.52) (N-ACETYL-BETA-D-GLUCOSAMINIDASE)
DE (BETA-N-ACETYLHEXOSAMINIDASE) (CHITOBINASE).
GN HEX.
OS Vibrio vulnificus.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN (1)

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 27562;
RX MEDLINE=93342067; Pubmed=8341694;

RA Somerville C.C., Colwell R.R.;
RA "Sequence analysis of the beta-N-acetylhexosaminidase gene of Vibrio
RT vulnificus: evidence for a common evolutionary origin of
RT hexosaminidases.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6751-6755(1993).

CC -1- FUNCTION: HYDROLYSIS OF TERMINAL, NON-REDUCING N-ACETYL-
CC BETA-D-GLUCOSAMINE RESIDUES IN CHITOBIOSE AND HIGHER ANNALOS, AND
CC IN GLYCOPROTEINS.

CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING N-ACETYL-
CC D-HEXOSAMINE RESIDUES IN N-ACETYL-BETA-D-HEXOSAMINIDES.
CC -1- PATHWAY: CONVERSION OF CHITIN TO GLCNAC.
CC -1- SIMILARITY: BELONGS TO FAMILY 20 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: L04544; AAA27527.1; -

DR HSSP: Q54468; IOBB.

DR InterPro: IPR001540; -

DR Pfam: PF00728; Glyco_hydro_20; 1.

DR PRINTS: PR00738; GLHYDRASE20.

KW Hydrolase: Glycosidase.

FT DISULFID 31 40 BY SIMILARITY.

FT DISULFID 377 385 BY SIMILARITY.

FT DISULFID 484 530 BY SIMILARITY.

FT ACT_SITE 519 519 CATALYTIC ACID (HINDS TO THE GLYCOSIDIC
LINKAGE) (BY SIMILARITY).

SQ SEQUENCE 847 AA; 94277 MW; 3C4405EFA34F3D14 CRC64;

Query Match 54.7%; Score 35; DB 1; Length 847;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12

Tue Jun 5 07:09:00 2001

Db 124 YVASEHLEGRN 135

Search completed: June 4, 2001, 12:24:05
Job time: 567 sec

us-09-284-107-24.rsp

Page 9

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:23:24 ; Search time 107.68 Seconds
(without alignments)
13.062 Million cell updates/sec

Title: US-09-284-107-24

Sequence: 1 FYMSKTLGRN 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.15:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	374	4	092663 homo sapien
2	64	100.0	375	4	092495 homo sapien
3	39	60.9	379	1	058234 pyrococcus
4	38	59.4	276	1	058131 pyrococcus
5	37	57.8	318	10	091PL9 arabidopsis
6	37	57.8	349	6	09M2T0 bos taurus
7	37	57.8	603	11	0922C5 mus musculi
8	37	57.8	1163	5	026654 sarcoptera
9	37	57.8	4005	4	013744 homo sapien
10	36	56.2	280	5	077335 plasmodium
11	36	56.2	351	2	047307 escherichia
12	36	56.2	415	10	091NV4 arabidopsis
13	36	56.2	669	5	09VRH9 drosophila
14	36	56.2	671	5	09VRH8 drosophila
15	36	56.2	813	10	043546 liliium long
16	36	56.2	1840	11	061818 mus musculi
17	36	56.2	5388	5	09UID0 leishmania
18	35	54.7	302	10	091UY1 arabidopsis
19	35	54.7	404	2	09KXC4 escherichia

20	35	54.7	404	9	09XJX3 bacterioph
21	35	54.7	464	2	09RH05 varliovora
22	35	54.7	549	2	086955 thermotoga
23	35	54.7	726	2	09WYF3 thermotoga
24	35	54.7	882	10	091UW0 arabidopsis
25	35	54.7	3265	5	09NMW1 leishmania
26	35	54.7	265	2	046059 coynnebacte
27	34	53.1	286	5	021741 caenorhabditi
28	34	53.1	387	2	069098 clostridium
29	34	53.1	452	14	098189 molluscum c
30	34	53.1	482	2	086912 sphingomonas
31	34	53.1	462	2	068616 synechococc
32	34	53.1	527	5	027471 caenorhabditi
33	34	53.1	529	4	09N216 homo sapien
34	34	53.1	531	2	053524 synechocyst
35	34	53.1	557	4	015059 homo sapien
36	34	53.1	643	3	012062 aspergillus
37	34	53.1	700	5	09N209 caenorhabditi
38	34	53.1	771	2	09WR09 thermotoga
39	34	53.1	1181	2	09SGP1 clostridium
40	34	53.1	1410	2	09KGT0 mycoplasma
41	34	53.1	1427	2	09KGX8 mycoplasma
42	33	51.6	107	5	021135 caenorhabditi
43	33	51.6	142	2	09KZ19 streptomyces
44	33	51.6	165	2	067921 aquifex aeo
45	33	51.6	167	5	018607 caenorhabditi

ALIGNMENTS

RESULT 1
ID 092663 PRELIMINARY; PRT: 374 AA.
AC 092663:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE FC GAMMA RECEPTOR I.
GN AL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=93055454; PubMed=1430234;
RA Porjes A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RT "Novel Fc gamma receptor I family gene products in human mononuclear
cells.";
RL J. Clin. Invest. 90:2102-2109(1992).
DR EMBL: L03418; AAA36049.1;
DR HSSP: P12319; 1ALT.
DR INTERPRO: IPR003006;
DR PIRAM: PF00047; 1g; 3.
DR PRODOM: PD002534; 1.
SO SEQUENCE 374 AA; 42632 MW; D33D59398CEEA699 CRC64;

Query Match Score 64; DB 4; Length 374;
Best local similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYMSKTLGRN 12
DB 229 FYMSKTLGRN 240
RESULT 2
ID 092495 PRELIMINARY; PRT: 375 AA.

AC 092495;
 DT 01-FEB-1997 (TREMBLrel..02, Created)
 DT 01-FEB-1997 (TREMBLrel..02, last sequence update)
 DT 01-OCT-2000 (TREMBLrel..15, last annotation update)
 DE FC GAMMA RECEPTOR TYPE I (FC GAMMA RIB-FC GAMMA RECEPTOR)
 GN CD64 OR FC<GAMMA>RIB
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE=93018827; PubMed=1402657;
 RA Benech P.D., Sasstry K.N., Iyer R.R., Eichbaum Q.G., Ravesh D.P.,
 RA Ezekowitz R.A.;
 RT "Definition of Interferon gamma-response elements in a novel human Fc
 RT gamma receptor gene (Fc gamma RIIb) and characterization of the gene
 RT structure."
 RL J. Exp. Med. 176:1115-1123(1992).
 DR EMBL; M91555; AA58414.1; JOINED
 DR EMBL; M91550; AA58414.1; JOINED
 DR EMBL; M91551; AA58414.1; JOINED
 DR EMBL; M91552; AA58414.1; JOINED
 DR EMBL; M91553; AA58414.1; JOINED
 DR EMBL; M91554; AA58414.1; JOINED
 DR EMBL; M91559; AAD13842.1; JOINED
 DR EMBL; M91570; AAD13842.1; JOINED
 DR EMBL; S45707; AAD13842.1; JOINED
 DR EMBL; S45708; AAD13842.1; JOINED
 DR EMBL; S45709; AAD13842.1; JOINED
 DR EMBL; S45704; AAD13842.1; JOINED
 DR EMBL; S45705; AAD13842.1; JOINED
 DR HSSP; P12319; IALT.
 DR INTERPRO; IPR003006;
 DR PFM; PFO0047; 19; 3.
 DR PRODOM; PD002534; -; 1.
 SO SEQUENCE 375 AA; 42881 MW; A84D464CT0DD0F91 CRC64;

Query Match 100.0%; Score 64; DB 4; Length 375;
 Best Local Similarity 100.0%; Pred. No. 0.00027;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
 DB 230 FYMGSKTLGRN 241

RESULT 3
 ID 058234 PRELIMINARY; PRT: 379 AA.
 AC 058234;
 DT 01-AUG-1998 (TREMBLrel..07, Created)
 DT 01-AUG-1998 (TREMBLrel..07, last sequence update)
 DT 01-JUN-2000 (TREMBLrel..14, last annotation update)
 DE HYPOTHEICAL 44.0 KDA PROTEIN PH0498.
 GN PH0498.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 NCBI_TaxID=53953;
 RX MEDLINE=96344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hainawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
 RL DNA Res. 5:55-76(1998).
 DR EMBL; AP000002; BAA29586.1;
 DR INTERPRO; IPR001072;
 DR PFM; PFO2003; DUF135; 1.
 DR PRODOM; PD014265; -; 1.
 KW Hypothetical protein.
 SO SEQUENCE 379 AA; 43959 MW; 195422BA5641E29 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 379;
 Best Local Similarity 58.3%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 12
 DB 273 FYIAERKLGRN 284

RESULT 4
 ID 058131 PRELIMINARY; PRT: 276 AA.
 AC 058131;
 DT 01-AUG-1998 (TREMBLrel..07, Created)
 DT 01-AUG-1998 (TREMBLrel..07, last sequence update)
 DT 01-JUN-2000 (TREMBLrel..14, last annotation update)
 DE HYPOTHEICAL 31.3 KDA PROTEIN PH0394.
 GN PH0394.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 NCBI_TaxID=53953;
 RX MEDLINE=96344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hainawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
 RL DNA Res. 5:55-76(1998).
 DR EMBL; AP000002; BAA29469.1;
 KW Hypothetical protein.
 SO SEQUENCE 276 AA; 31330 MW; 535BEDAF500983B CRC64;

Query Match 59.4%; Score 38; DB 1; Length 276;
 Best Local Similarity 70.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 10
 DB 264 FYAGSKVFRG 273

RESULT 5
 ID 091PL9 PRELIMINARY; PRT: 318 AA.
 AC 091PL9;
 DT 01-OCT-2000 (TREMBLrel..15, Created)
 DT 01-OCT-2000 (TREMBLrel..15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel..15, last annotation update)
 DE F2J10.16 PROTEIN.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
 RP [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Sakano H., Liu S.X., Yu G., Lee J., Lenz C., Pham P., Toriumi M.,
 RA Chin C., Chlou J., Choi E., Chung M., Gonzalez A., Howng B., Liu A.,
 RA Vaysberg M., Altati H., Brooks S., Buehler E., Chao Q., Conn L.,
 RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
 RA Miranda M., Nguyen M., Palm C., Shinn P., Southwick A., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT The sequence of BAC F2J10 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC015445; AAF76448.1;
 SO SEQUENCE 318 AA; 35231 MW; E722F2FBED9FD961 CRC64;

Query Match 57.8%; Score 37; DB 10; Length 318;
 Best Local Similarity 87.5%; Pred. No. 37;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 SKTLGRN 12
 DB 285 SKTLGRN 292

RESULT 6
 ID 09M2T0 PRELIMINARY; PRT: 349 AA.
 AC 09M2T0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR I.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yan Y., Li X., Wang A., Zhang G.;
 RT "Molecular cloning and identification of full-length cDNA encoding
 RT high affinity Fc receptor for bovine Igg (Fc gamma RI).";
 RL Vet. Immunol. Immunopathol. 75:151-159(2000).
 DR EMBL: AF162866; AAF80477.1;
 KW Receptor.
 SO SEQUENCE 349 AA; 39608 MW; D0B7B2EF9408C02 CRC64;

Query Match 57.8%; Score 37; DB 6; Length 349;
 Best Local Similarity 63.6%; Pred. No. 41;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FYMGSKTLGRN 11
 DB 229 FYMGSKTLGRN 239

RESULT 7
 ID 09Z2C5 PRELIMINARY; PRT: 603 AA.
 AC 09Z2C5;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE MYOFIBRILARIN (EC 3.1.3.48).
 GN MTM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-98409499; PubMed-9736772;
 RA Laporte J., Blondeau F., Buj-Bello A., Tentler D., Kretz C., Dahl N.,
 RA Mandel J.L.;
 RT "Characterization of the myotubularin dual specificity phosphatase
 RT gene family from yeast to human.";
 RL Hum. Mol. Genet. 7:1703-1712(1998).
 DR EMBL: AF073996; AAC77821.1;
 DR INTERPRO: IPR000387;
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 1.
 KW Hydrolase.
 SO SEQUENCE 603 AA; 69594 MW; 65929312F73FE2EC7 CRC64;

Query Match 57.8%; Score 37; DB 11; Length 603;
 Best Local Similarity 63.6%; Pred. No. 74;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 YMGSKTLGRN 12
 DB 95 YMGSKTLGRN 105

RESULT 8
 ID 026654 PRELIMINARY; PRT: 1163 AA.
 AC 026654;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE STORAGE PROTEIN-BINDING PROTEIN.
 OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Oestroidea; Sarcophagidae; Sarcophaga.
 OX NCBI_TaxID=7386;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95181458; PubMed-7876233;
 RA Chung S., Kubo T., Natoli S.;
 RT "Molecular cloning and sequencing of arylphorin-binding protein in
 RT protein granules of the Sarcophaga fat body. Implications of a post-
 RT translational processing mechanism.";
 RL J. Biol. Chem. 270:4624-4631(1995).
 DR EMBL: D29741; BAA06161.1;
 DR HSSP: P04253; IOXY.
 DR INTERPRO: IPR000896;
 DR PFAM: PF00372; hemocyanin; 2.
 SO SEQUENCE 1163 AA; 133401 MW; D357B901ED3E0416 CRC64;

Query Match 57.8%; Score 37; DB 5; Length 1163;
 Best Local Similarity 54.5%; Pred. No. 1.5e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 YMGSKTLGRN 12
 DB 464 YMGSKTLGRN 474

RESULT 9
 ID 013744 PRELIMINARY; PRT: 4005 AA.
 AC 013744;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE ALL-1 PROTEIN.
 GN ALL-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96290553; PubMed=8703835;
 RA Nilsson I., Loechner K., Stiegler G., Grell J., Beck J.D., Fey G.H.,
 RA Marchalek R.;
 RT "Exon/Intron structure of the human ALL-1 (MLL) gene involved in
 RT translocations to chromosomal region 11q23 and acute leukaemias.";
 RL Br. J. Haematol. 93:966-972(1996);
 [2]
 RP SEQUENCE OF 1245-1639 FROM N.A.
 RX MEDLINE=95315013; PubMed=7794749;
 RA Marchalek R., Grell J.D., Fey G.H.,
 RA Zwickelmeier I., Beck J.D., Nilsson I., Stiegler G.,
 RT "Molecular analysis of the chromosomal breakpoint and fusion
 RT transcripts in the acute lymphoblastic stem cell line with chromosomal
 RT translocation t(4;11).";
 RL Br. J. Haematol. 90:308-320(1995);
 DR EMBL: 269745; CAA93625.1; JOINED
 DR EMBL: 269746; CAA93625.1; JOINED
 DR EMBL: 269747; CAA93625.1; JOINED
 DR EMBL: 269748; CAA93625.1; JOINED
 DR EMBL: 269749; CAA93625.1; JOINED
 DR EMBL: 269750; CAA93625.1; JOINED
 DR EMBL: 269751; CAA93625.1; JOINED
 DR EMBL: 269752; CAA93625.1; JOINED
 DR EMBL: 269753; CAA93625.1; JOINED
 DR EMBL: 269754; CAA93625.1; JOINED
 DR EMBL: 269755; CAA93625.1; JOINED
 DR EMBL: 269756; CAA93625.1; JOINED
 DR EMBL: 269757; CAA93625.1; JOINED
 DR EMBL: 269758; CAA93625.1; JOINED
 DR EMBL: 269759; CAA93625.1; JOINED
 DR EMBL: 269760; CAA93625.1; JOINED
 DR EMBL: 269761; CAA93625.1; JOINED
 DR EMBL: 269762; CAA93625.1; JOINED
 DR EMBL: 269763; CAA93625.1; JOINED
 DR EMBL: 269764; CAA93625.1; JOINED
 DR EMBL: 269765; CAA93625.1; JOINED
 DR EMBL: 269766; CAA93625.1; JOINED
 DR EMBL: 269767; CAA93625.1; JOINED
 DR EMBL: 269768; CAA93625.1; JOINED
 DR EMBL: 269769; CAA93625.1; JOINED
 DR EMBL: 269770; CAA93625.1; JOINED
 DR EMBL: 269772; CAA93625.1; JOINED
 DR EMBL: 269773; CAA93625.1; JOINED
 DR EMBL: 269774; CAA93625.1; JOINED
 DR EMBL: 269775; CAA93625.1; JOINED
 DR EMBL: 269776; CAA93625.1; JOINED
 DR EMBL: 269777; CAA93625.1; JOINED
 DR EMBL: 269778; CAA93625.1; JOINED
 DR EMBL: 269779; CAA93625.1; JOINED
 DR EMBL: 269780; CAA93625.1; JOINED
 DR EMBL: X83604; CAA8584.1;
 DR INTERPRO: IPR000637;
 DR INTERPRO: IPR001214;
 DR INTERPRO: IPR001487;
 DR INTERPRO: IPR001965;
 DR INTERPRO: IPR002048;
 DR INTERPRO: IPR002857;
 DR PFAM: PF00628; PHD; 3;
 DR PFAM: PF00856; SET; 1;
 DR PFAM: PF02008; zf-CXXC; 1;
 DR PROSITE: PS00018; EF_HAND; UNKNOWN; 1;
 DR PROSITE: PS00354; HMG_LY; UNKNOWN; 1;
 DR PROSITE: PS50014; BROMODOMAIN; 2;
 SQ SEQUENCE 4005 AA; 435828 MW; F2FB410757B77313 CRC64;

Query Match 57.8%; Score 37; DB 4; Length 4005;
 Best Local Similarity 66.7%; Pred. No. 5.6e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12
 Db 872 FYPGQTERGRN 883
 RESULT 10
 077335 PRELIMINARY; PRT; 280 AA.
 AC 077335;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE PFC0410W PROTEIN.
 GN PFC0410W.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mungall K., Lawson D., Barrell B.;
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: 298347; CAB1120.2;
 DR INTERPRO: IPR001064;
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN; 1;
 SQ SEQUENCE 280 AA; 33089 MW; 1C37A2BCD82A846 CRC64;

Query Match 56.2%; Score 36; DB 5; Length 280;
 Best Local Similarity 50.0%; Pred. No. 50;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FYMGSKTLGRN 12
 Db 98 FYDNKVFGRKN 109
 RESULT 11
 047307 PRELIMINARY; PRT; 351 AA.
 AC 047307;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE TRANSPOSASE.
 GN TNP1294.
 OS Escherichia coli.
 OC Plasmid pUD2380.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Brynys B.;
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON-IS1294;
 RA Albiger B., Bennett P.M., Commanducci A., Dodd H.M., Lett M.C.;
 RT "pUB2380: a cold-like resistance plasmid with three replication
 RT origins for three different functions.";
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ008006; CAA07835.1;
 KW Plasmid.
 SQ SEQUENCE 351 AA; 40283 MW; 46FAD7C01E075CAD CRC64;

Query Match 56.2%; Score 36; DB 2; Length 351;
 Best Local Similarity 63.6%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 YMGSKTLGRN 12
 Db 11 11 11

DB 237 YMSKKTAGRN 247

RESULT 12

ID 09LNV4 PRELIMINARY; PRT: 415 AA.

AC 09LNV4

DT 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)

DE F2265.31.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

[1]

RP SEQUENCE FROM N.A.

RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bel O., Chin C., Chou J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharisky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaverl A., Tortumli M., Vaysberg M., Yu G., Federspiel N.A., Theologis A., Ecker J.R.,

RT Genomic sequence for Arabidopsis thaliana BAC F2265 from chromosome 1.

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Ecker J.R.,

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Ecker J.R.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA Ecker J.R.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bel O., Chin C., Chou J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharisky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaverl A., Tortumli M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A., Ecker J.,

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC022464; AAF79558.1;

SQ SEQUENCE 415 AA; 48371 MW; 8C86C784B0AC2C68 CRC64;

QY 2 YMSKKTAGRN 12

DB 199 YLGRTARGRN 209

Query Match 56.2%; Score 36; DB 10; Length 415;

Best Local Similarity 54.5%; Pred. No. 77;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 13

ID 09VRH9 PRELIMINARY; PRT: 669 AA.

AC 09VRH9

DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)

DE SIGA PROTEIN.

GN SIGA

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN-BREKLEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Cealiker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amandiles P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G., Abril J.F., Adayani A., An H.-D., Andrews-Planckoch C., Baldwin D., Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottler P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., DePamphilis M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyama C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshell A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T., Slier E., Spradling A.C., Stapleton M., Strong R., Sun E., Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaverl J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

DR EMBL: AEO03568; AAF50819.1;

DR EMBASE; FBgn0003423; SIGA.

DR INTERPRO; IPR002872;

DR PFAM; PF01619; Pro_dh; 1.

SQ SEQUENCE 669 AA; 75951 MW; 89C68093DF0D828E CRC64;

QY 1 FYMSKKTARG 10

DB 491 FYMSKKTARG 500

Query Match 56.2%; Score 36; DB 5; Length 669;

Best Local Similarity 60.0%; Pred. No. 13;+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 14

ID 09VRH8 PRELIMINARY; PRT: 671 AA.

AC 09VRH8

DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)

DE SIGA PROTEIN.

GN SIGA

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Makos G.L.G.,
 RA Abril J.F., Abghyani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktiroglu U., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gariellean A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Giodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos A.C., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling I.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Weissbach J.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003568; AAF50820.1;
 DR FLYBASE: FBgn003423; slga.
 DR INTERPRO: IPR002872;
 DR PFAM: PF01619; Pro_dh; 1;
 SQ SEQUENCE 671 AA; 76113 MW; 62FA4370823FEF16 CRC64;

Query Match 56.2%; Score 36; DB 5; Length 671;
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYGSKTLRG 10
 |||:|:|:
 DB 493 FYFGAKLVNG 502

RESULT 15
 Q43546 PRELIMINARY; PRT; 813 AA.
 AC Q43546;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE SERINE PROTEINASE (FRAGMENT).
 GN LIM9.
 OS Liliium longiflorum (Trumpet lily).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Liliales; Liliaceae; Liliium.

OX NCBI_TaxID=4690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, HINOMOTO; TISSUE=FLORAL BUD;
 RA Tabata S.;
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, HINOMOTO; TISSUE=FLORAL BUD;
 RX MEDLINE=96051386; PubMed=7584025;
 RA Kobayashi T., Kobayashi E., Sato S., Hotta Y., Miyazima N., Tanaka A.,
 RA Tabata S.;
 RT "Characterization of cDNAs induced in meiotic prophase in lily
 RT microsporocytes.";
 RL DNA Res. 1:15-26(1994).
 DR EMBL: D21815; BA004839.1;
 DR HSSP: P00782; 2SPT.
 DR MEROPS: S08.092;
 DR MENDEL: 11876; L110:1086;11876.
 DR INTERPRO: IPR00209;
 DR PFAM: PF00082; Peptidase_S8; 2.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 FT NON_TER 1
 SQ SEQUENCE 813 AA; 86372 MW; ECFE88911C4FE6E3 CRC64;

Query Match 56.2%; Score 36; DB 10; Length 813;
 Best Local Similarity 63.6%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 YGSKTLRG 12
 |||||:
 DB 505 YGSKTLRG 515

Search completed: June 4, 2001, 12:23:26
 Job time: 589 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:29 ; Search time 106.14 Seconds
(without alignments)
6.463 Million cell updates/sec

Title: US-09-284-107-25

Perfect score: 57

Sequence: 1 TSSEXQILTRAR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq.0401:*

1:	/SID56/gcgdata/geneseq/AA1980.DAT:*
2:	/SID56/gcgdata/geneseq/AA1981.DAT:*
3:	/SID56/gcgdata/geneseq/AA1982.DAT:*
4:	/SID56/gcgdata/geneseq/AA1983.DAT:*
5:	/SID56/gcgdata/geneseq/AA1984.DAT:*
6:	/SID56/gcgdata/geneseq/AA1985.DAT:*
7:	/SID56/gcgdata/geneseq/AA1986.DAT:*
8:	/SID56/gcgdata/geneseq/AA1987.DAT:*
9:	/SID56/gcgdata/geneseq/AA1988.DAT:*
10:	/SID56/gcgdata/geneseq/AA1989.DAT:*
11:	/SID56/gcgdata/geneseq/AA1990.DAT:*
12:	/SID56/gcgdata/geneseq/AA1991.DAT:*
13:	/SID56/gcgdata/geneseq/AA1992.DAT:*
14:	/SID56/gcgdata/geneseq/AA1993.DAT:*
15:	/SID56/gcgdata/geneseq/AA1994.DAT:*
16:	/SID56/gcgdata/geneseq/AA1995.DAT:*
17:	/SID56/gcgdata/geneseq/AA1996.DAT:*
18:	/SID56/gcgdata/geneseq/AA1997.DAT:*
19:	/SID56/gcgdata/geneseq/AA1998.DAT:*
20:	/SID56/gcgdata/geneseq/AA1999.DAT:*
21:	/SID56/gcgdata/geneseq/AA2000.DAT:*
22:	/SID56/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	12	19 W60557	Oligopeptide from
2	57	100.0	261	20 W33183	Human sfc-gammari
3	57	100.0	344	13 R22549	Human macrophage-s
4	57	100.0	344	17 R91439	Human FCRI (CDNA c
5	57	100.0	344	21 Y96183	Human macrophage-s
6	57	100.0	374	13 R20811	Human macrophage-s
7	57	100.0	374	13 R22550	Human macrophage-s
8	57	100.0	374	17 R91438	Human FCRI (CDNA c
9	57	100.0	374	17 W00859	Human FCRI (CDNA c
10	57	100.0	374	19 W80448	Human Fc receptor
11	57	100.0	374	19 W97833	Human Fc receptor

ALIGNMENTS

12	57	100.0	374	19 W97834	Human Fc receptor
13	57	100.0	374	21 Y96134	Human macrophage-s
14	57	100.0	374	21 Y96226	Human high affinity
15	57	100.0	377	20 W86195	Human Fc receptor
16	57	100.0	399	21 B43683	Human cancer assoc
17	36	63.2	410	12 R12428	Hydrid Fc(gamma)RI
18	36	63.2	1879	14 R42455	Enzyme involved in
19	36	63.2	2004	17 R99464	Biosynthetic enzym
20	36	63.2	2004	19 W37052	S. putrefaciens Ep
21	36	63.2	2004	20 W89402	S. putrefaciens PK
22	36	63.2	2004	21 B10469	Shewanella putrefa
23	34	59.6	88	20 Y35386	Amino acid sequenc
24	33	57.9	193	18 W55378	H. pylori ORF 07CP
25	33	57.9	353	21 Y67867	Mouse Mgbp5 G-prot
26	33	57.9	371	21 B54247	Human pancreatic c
27	33	57.9	395	20 Y31732	Human cell cycle r
28	33	57.9	696	18 W55698	H. pylori ORF 14gp
29	33	57.9	696	20 Y17207	H. pylori outer me
30	33	57.9	749	8 P70286	Protein encoded by
31	33	57.9	1165	14 R37309	Cardiac adenylyl c
32	32	56.1	25	20 Y25887	Human secreted pro
33	32	56.1	283	21 G39208	Arabidopsis thalia
34	32	56.1	284	21 G39207	Arabidopsis thalia
35	32	56.1	428	18 W11883	Arabidopsis thalia
36	32	56.1	451	21 G39206	Arabidopsis thalia
37	32	56.1	506	21 Y94986	Human secreted pro
38	32	56.1	525	20 Y25881	Human secreted pro
39	32	56.1	623	15 R49578	Sequence of bovine
40	32	56.1	802	21 G39222	Arabidopsis thalia
41	32	56.1	969	21 G39221	Arabidopsis thalia
42	32	56.1	992	21 G39220	Arabidopsis thalia
43	31	54.4	58	21 G00321	Human secreted pro
44	31	54.4	185	21 Y58587	Sorangium cellulos
45	31	54.4	235	21 Y74710	Neisseria gonorrhoe

RESULT 1
ID W60557
XX W60557 standard; peptide: 12 AA.
XX
AC W60557:
XX
DT 18-AUG-1998 (first entry)
XX
DE Oligopeptide from extracellular domain of CD64.
XX
KW Extracellular domain; CD64; identification; antibody;
KW Immunohistochemical; immunofluorescent analysis; detection;
KW cell transformation; mutation; anti; oncogene.
XX
OS Synthetic.
XX
XX W09815833-A1.
PN
PD 16-APR-1998.
XX
XX 07-OCT-1997; 97WO-NL00557.
PF
XX 08-OCT-1996; 96EP-0202791.
PR
XX (UYDT-) RIKUSUNIV UTRICHT.
PA De Kruif CA, Logtenberg T;
XX WPI; 1998-240964/21.
XX
XX Identifying peptide(s) binding specifically to protein target - by
PT expressing on phage surface and testing for binding to immobilised
PT oligopeptide derived from the target, useful for, e.g. identifying
PT specific antibodies

XX Example 1; Page 29; 40pp; English.

XX Synthetic oligopeptides W60537-61 are derived from the extracellular
CC domain of CD64. They were synthesised on 25 polyethylene rods as
CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
CC affinity binding to the CD64-covered rods. The specification describes a
CC method for the identification of a peptide able to bind specifically to a
CC target protein. The method comprises displaying the peptide on the
CC surface of a replicable display package, synthesising oligopeptides
CC derived from the target protein on a solid phase, and testing for binding
CC between the peptide and oligopeptides. The method is used to screen large
CC peptide libraries, especially to detect antibodies, or their fragments,
CC that bind to cell markers or that can differentiate between different
CC forms of the same protein, including bispecific antibodies that bind to
CC two non-overlapping epitopes on the same monomeric antigen or two
CC epitopes on different molecules. The genes/oligonucleotides that encode
CC selected peptides can be isolated and used for recombinant production of
CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
CC immunofluorescent analysis, and also to detect cell transformation caused
CC by mutation in (anti)oncogenes.

SO Sequence 12 AA;

Query Match 100.0%; Score 57; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYQILTARR 12
| | | | | | | | | | | | | |
DB 1 tsseyqiltarr 12

RESULT 2
Y33183
ID Y33183 standard; Protein; 261 AA.

AC Y33183;

DT 15-NOV-1999 (first entry)

DE Human sFc-gammamRI protein fragment.

XX Fc receptor; Fc-gammamRIa; human; FCR; model; three-dimension; 3-D;
KW atomic coordinate; bioactive compound design; computer-assisted;
KW drug design; therapy; inhibitor; Fc-gammamRI; Fc-epsilonRI; IgG; IgE;
KW tissue damage; hypersensitivity; inflammatory cell recruitment;
KW inflammatory modulator; Fc-gammamRIa; immune function regulation;
KW anti-inflammatory; immunoprotective; sFc-gammamRI.

OS Homo sapiens.

PN W09940117-A1.

PD 12-AUG-1999.

PF 04-FEB-1999; 99WO-IB00367.

PR 11-SEP-1998; 98US-0099994.

PR 06-FEB-1998; 98US-0073972.

PA (ILEX-) ILEXUS PTY LTD.

PI Baeil JB, Epa V, Garrett TPJ, Hogarth PM, Matthews BR;

PI Maxwell KF, McCarthy JD, McKenzie JFC, Pietersz GA;

PI Powell MS;

DR WPI: 1999-539978/45.
XX Three-dimensional structures and models of Fc receptors, useful in
PT computer-assisted drug design

PS Claim 11; Page 316-317; 326pp; English.

XX This invention describes a novel model of an Fc receptor (FCR) protein
CC representing a three-dimensional (3-D) structure that substantially
CC conforms to the specified atomic coordinates. Computer model images of
CC the FCR can be used to design bioactive chemical compounds, e.g.,
CC oligonucleotides, peptides, peptidomimetics and small organic molecules,
CC by computer-assisted methods of drug design. Therapeutic compositions
CC that inhibit the activity of Fc-gammamRI or Fc-epsilonRI can be used to
CC reduce IgG- or IgE-mediated tissue damage, e.g. IgG- or IgE-mediated
CC hypersensitivity, recruitment of inflammatory cells or release of
CC inflammatory modulators. The therapeutic compositions can also be used to
CC treat disease in animals. The knowledge of the 3-D structure of Fc-gamma
CC RIA and models of other FCR provides a means for designing and producing
CC compounds that regulate immune function and inflammation in an animal,
CC including humans (i.e. structure based drug design). For example,
CC chemical compounds can be designed to block binding of immunoglobulin to
CC an Fc receptor protein using various computer programs and models. The
CC products of the invention have anti-inflammatory and immunoprotective
CC activity. This sequence represents the human sFc-gammamRI protein
CC fragment described in the method of the invention.

SO Sequence 261 AA;

Query Match 100.0%; Score 57; DB 20; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYQILTARR 12
| | | | | | | | | | | | | |
DB 222 tsseyqiltarr 233

RESULT 3
R22549
ID R22549 standard; Protein; 344 AA.

AC R22549;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FcRI receptor encoded by clone p98.

XX Rapid immunoselection cloning technique; cell surface antigen;
KW immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN W09201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GEO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amiot M;

PI WPI: 1992-056864/07.

PI N-PSDB; Q21179.

PT New CD53 cell surface antigen and DNA encoding it - for
PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

XX This amino acid sequence was predicted from the cDNA sequence of
CC cDNA clone p98/X2. It differs from the sequence predicted from
CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
CC while clone p98/X2 predicts a Thr residue. At position 58, p90

CC (see 021180) predicts Val and p98/X2 predicts Leu. Sequences
 CC predicted from all 3 clones show the typical features of a type I
 CC integral membrane protein and include a short hydrophobic signal
 CC sequence, a single 21-residue hydrophobic membrane-spanning domain,
 CC and a short, highly charged cytoplasmic domain. The extracellular
 CC portion contains six potential N-linked glycosylation sites and six
 CC Cys residues distributed among three C2 set Ig-related domains.
 CC A fusion protein of FCRI and a receptor ligand will be helpful to
 CC increase the potency of antibodies in therapy.
 SO Sequence 344 AA:

Query Match 100.0%; Score 57; DB 13; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYOILTFARR 12
 DB 241 tsseyqillarr 252

RESULT 4
 ID R91439 standard; Protein; 344 AA.
 AC R91439;
 DT 30-OCT-1996 (first entry)
 DE Human FCRI (CDNA clone p98 product).
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KM therapy; diagnosis; vector; FCRI; Fc receptor.
 OS Homo sapiens.

Key Location/Qualifiers
 FT Misc-difference 58 /note- "amino acid 58 is Leu in clone p135
 translated product."

US5506126-A.
 09-APR-1996.
 25-FEB-1988; 88US-0160416.
 01-DEC-1992; 92US-0983647.
 25-FEB-1988; 88US-0160416.
 13-JUL-1989; 89US-0379076.
 13-JUL-1990; 90US-0553759.
 18-OCT-1993; 93US-0139273.
 (GENE) GEN HOSPITAL CORP.
 Aruffo A, Seed B;
 WPI; 1996-200279/20.
 N-PSDB; T14718.

Cloning of CDNA encoding cell surface antigen - useful for isolation
 of diagnostic and therapeutic proteins
 Example 10; Column. 55-56; 79pp; English.

The amino acid sequence (R91439) of human FCRI was detd. from a
 CC cDNA clone, p98 (T14718), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. 2 other cDNA clones, p135 (T14717) and p90 (T14719),
 CC coded for a variants of the FCRI sequence: the C-terminal sequence
 CC of the p98 product is truncated compared with those of the p135

CC and p90 products (see also R91438 and M00859). A fusion protein
 CC of FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.

SO Sequence 344 AA:

Query Match 100.0%; Score 57; DB 17; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYOILTFARR 12
 DB 241 tsseyqillarr 252

RESULT 5
 ID Y96183 standard; Protein; 344 AA.
 AC Y96183;
 DT 19-DEC-2000 (first entry)
 DE Human macrophage-specific FCRI.
 KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
 KM panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KM immune disorder; infection; asthma; immune-complex disease;
 KM amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.
 OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 15 /note- "encoded by GCG"

FT Misc-difference 38 /note- "encoded by ACC"

FT Misc-difference 50 /note- "encoded by CCG"

FT Misc-difference 51 /note- "encoded by ACC"

FT Misc-difference 55 /note- "encoded by CAC"

FT Misc-difference 56 /note- "encoded by TCC"

FT Misc-difference 60 /note- "encoded by CCC"

FT Misc-difference 64 /note- "encoded by CAG"

FT Misc-difference 82 /note- "encoded by CAA"

FT Misc-difference 116 /note- "encoded by GAA"

FT Misc-difference 117 /note- "encoded by GCA"

FT Misc-difference 122 /note- "encoded by TTC"

FT Misc-difference 123 /note- "encoded by ACC"

FT Misc-difference 126 /note- "encoded by CCG"

FT Misc-difference 129 /note- "encoded by CAT"

FT Misc-difference 134 /note- "encoded by AAT"

FT Misc-difference 136 /note- "encoded by GTT"

FT Misc-difference 139 /note- "encoded by CCA"

FT Misc-difference 140 /note- "encoded by AAT"

FT Misc-difference 213 /note- "encoded by AAT"

Sequence 344 AA:

Query Match	100.0%;	Score 57;	DB 21;	Length 344;
Best Local Similarity	100.0%;	Pred. No. 0.0011;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1 TSSEYQILTARR 12.			

```

Db      241 tsseyqlltarr 252

RESULT      6
ID          R20811
AC          R20811 standard; Protein; 374 AA.
XX
XX          R20811;
DE          21-MAY-1992 (first entry)
XX
XX          Human macrophage-specific FCRI receptor encoded by clone p135.
KW          Rapid immunoselection cloning technique; cell surface antigen;
XX          Immunodiagnosis; high affinity receptor.
OS          Homo sapiens.
XX
XX          WO9201049-A.
PN          23-JAN-1992.
XX
XX          15-JUL-1990; 90WO-US04986.
PF          13-JUL-1990; 90US-0553759.
XX
XX          (GEMO-) GEN HOSPITAL CORP.
PA          Seed B, Aruffo A, Amlot M;
PI          WPI; 1992-056864/07.
DR          N-PSDB; Q21178.
XX
XX          New CD53 cell surface antigen and DNA encoding it - for
PT          immuno-therapy and diagnosis of hematopoietic neoplasms, etc.
XX
XX          Example 10; Page 94a; 160pp; English.
XX
CC          This amino acid sequence was predicted from the cDNA sequence of
CC          cDNA clone p135. It differs from the sequence predicted from
CC          clones p98/X2 and p90 (see Q21179 and Q21180, respectively) at
CC          position 25: p135 encodes a Ser residue and the other two clones
CC          predict a Thr residue. At position 58, p135 predicts Leu and p90
CC          predicts Val. Sequences predicted from all 3 clones show the
CC          typical features of a type I integral membrane protein and include
CC          a short hydrophobic signal sequence, a single 21-residue
CC          cytoplasmic domain. The extracellular portion contains six
CC          potential N-linked glycosylation sites and six Cys residues
CC          distributed among three C2 set Ig-related domains. A fusion protein
CC          of FcR1 and a receptor ligand will be helpful to increase the
CC          potency of antibodies in therapy.
XX
XX          Sequence      374 AA;
SQ

```

```

DE Human macrophage-specific FCRI receptor encoded by clone p90.
XX
KW Rapid immunoselection cloning technique; cell surface antigen;
KM immunodiagnosis; high affinity receptor.
XX
OS Homo sapiens.
XX
PN M09201049-A.
XX
PD 23-JAN-1992.
XX
PE 15-JUL-1990; 90WO-US04986.
XX
PR 13-JUL-1990; 90US-0553759.
XX
PA (GEHO-) GEN HOSPITAL CORP.
XX
PI Seed B., Aruffo A., Amiot M;
XX
DR WPI. 1992-056864/07.
XX
N-PSDB; Q21180.
XX
PT New CD53 cell surface antigen and DNA encoding it - for
XX immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
PS Example 10; Page 94a; 160pp; English.
XX
CC This amino acid sequence was predicted from the cDNA sequence of
CC cDNA clone p90. It differs from the sequence predicted from
CC clone p135 (see Q21178) at position 25: p135 encodes a Ser residue
CC and clone p90 predicts a Thr residue. At position 58, p90 predicts
CC Val while clones p135 and p98/X2 (see Q21179) both predict Leu.
CC Sequences predicted from all 3 clones show the typical features of
CC a type I integral membrane protein and include a short hydrophobic
CC signal sequence, a single 21-residue hydrophobic membrane-spanning
CC domain, and a short, highly charged cytoplasmic domain. The
CC extracellular portion contains six potential N-linked glycosylation
CC sites and six Cys residues distributed among three C2 set Ig-related
CC domains. A fusion protein of FCRI and a receptor ligand will be
CC helpful to increase the potency of antibodies in therapy.
CC
SQ Sequence 374 AA;

Query Match 100.0%; Score 57; DB 13; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 TSSEYOITARR 12
   |||||
Db 241 tsseyqiltarr 252

RESULT 8
R91438 R91438 standard; Protein; 374 AA.
XX
AC AC
XX
R91438;
XX
DT DT
XX 30-OCT-1996 (first entry)
XX
DE Human FCRI (cDNA clone p135 product).
XX
KW Cell surface antigen; cloning; immunoselection; immunotherapy;
KW therapy; diagnosis; vector; FCRI; Fc receptor.
XX
OS Homo sapiens.
XX
FH FH Location/Qualifiers
FT FT Misc-difference 25
FT /note= "amino acid 25 is Thr in clone p90 and p98
FT translated products"
FT Misc-difference 58

```

FT		/note= "amino acid 58 is Val in p90-clone
FT		translated product"
XX		
FN	US5506126-A.	
PD	09-APR-1996.	
XX		
PF	25-FEB-1988;	88US-0160416.
XX		
PR	01-DEC-1992;	92US-0983647.
PR	25-FEB-1988;	88US-0160416.
PR	13-JUL-1989;	89US-0379076.
PR	13-JUL-1990;	90US-0553759.
PR	18-OCT-1993;	93US-0139273.
XX		
PA	(GENO) GEN HOSPITAL CORP.	
PI	Aruifo A, Seed B;	
XX		
DR	WPI; 1996-200279/20.	
DR	N-PSDB; T14717.	
PT	Cloning of cDNA encoding cell surface antigen - useful for isolation	
PT	of diagnostic and therapeutic proteins	
XX		
PS	Example 10; Column 55-56; 79pp; English.	
XX		
CC	The amino acid sequence (R91438) of human FCRI was detd. from a	
CC	cDNA clone, p135 (T14717), obcd. from a cDNA library using an	
CC	immunoselection cloning method. FCRI is a high affinity receptor	
CC	for the Fc portion of IgG, normally located on cell surfaces of	
CC	macrophages. Another isolated cDNA clone, p90 (T14719), coded	
CC	for a variant (W00859) of the FCRI sequence, and a third clone, p98	
CC	(T14718), coded for an FCRI (R91439) having a different C-terminal	
CC	sequence. A fusion protein of FCRI and a receptor ligand will	
CC	be helpful in increasing the potency of antibodies in therapy.	
XX		
SQ	Sequence 374 AA:	
	Query Match .100.0%; Score 57; DB 17; Length 374;	
	Best Local Similarity 100.0%; Pred. No. 0.0012;	
	Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 TSSEYOILFARR 12	
DB	241 tsseygiltarr 252	
RESULT 9		
ID W00859	standard; Protein; 374 AA.	
AC W00859;		
DT 30-OCT-1996	(first entry)	
DE Human FCRI (cDNA clone p90 product).		
XX Cell surface antigen; cloning; immunoselection; immunotherapy;		
KM Therapy; diagnosis; vector; FCRI; Fc receptor.		
OS Homo sapiens.		
XX US5506126-A.		
PV 09-APR-1996.		
XX 25-FEB-1988;	88US-0160416.	
PF 01-DEC-1992;	92US-0983647.	
PR 25-FEB-1988;	88US-0160416.	
PR 13-JUL-1989;	89US-0379076.	

PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 (GCHO) GEN HOSPITAL CORP.
 PA Aruffo A, Seed B;
 PI WPI: 1996-200279/20.
 DR N-PSDB; T14719.
 XX Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 PS Example 10; Column 55-56; 79pp; English.
 XX The amino acid sequence (W00859) of human FCRI was detd. from a
 CC cDNA clone, p90 (T14719), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Other cDNA clones (see also T14717-18) coded for
 CC variants (R91438-39) of the FCRI sequence. A fusion protein of
 CC FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.
 CC Sequence 374 AA;
 SQ

Query Match 100.0%; Score 57; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYQILTRR 12
 |||||
 Db 241 tsseyqiltrr 252

RESULT 10
 W80448 ID W80448 standard; Protein; 374 AA.
 XX W80448;
 AC
 XX 07-JUN-1999 (first entry)
 DE Human Fc receptor I.
 XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 KM cloning.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 2 /note= "encoded by TGG"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by CTG"
 FT Misc-difference 60 /note= "encoded by CCC"
 FT Misc-difference 77 /note= "encoded by AAT"
 FT Misc-difference 85 /note= "encoded by TCC"
 FT Misc-difference 99 /note= "encoded by CAA"
 FT Misc-difference 103 /note= "encoded by CCC"
 FT Misc-difference 141 /note= "encoded by GGC"
 FT Misc-difference 159 /note= "encoded by GGC"

FT /note= "encoded by AAC"
 FT Misc-difference 171 /note= "encoded by ATG"
 FT Misc-difference 176 /note= "encoded by GTC"
 FT Misc-difference 256 /note= "encoded by GGC"
 FT
 XX US5830731-A.
 XX
 XX 03-NOV-1998.
 XX
 XX 21-MAY-1997; 97US-0861205.
 XX
 XX 01-DEC-1992; 92US-0983647.
 XX 25-FEB-1988; 88US-0160416.
 XX 13-JUL-1989; 89US-0379076.
 XX 13-JUL-1990; 90US-0553759.
 XX 21-MAY-1997; 97US-0861205.
 XX
 XX (GCHO) GEN HOSPITAL CORP.
 XX Aruffo A, Seed B;
 XX WPI: 1998-609251/51.
 XX N-PSDB; V63456.
 XX
 XX New cloning vector and polylinker - based on existing sequences for
 XX efficient cloning and expression of mammalian cDNA(s), especially
 XX human lymphocyte antigenic sequences
 XX
 XX Example 10; Column 53-54; 75pp; English.

This is the amino acid sequence of human Fc receptor I (FCRI), as
 deduced from cDNA clone p135 (see V63456) isolated using a rapid
 immunoselection cloning method from a cDNA library expressed in COS
 cells. The cDNA library was constructed from polyA RNA of cells
 from a single patient undergoing extracorporeal Interleukin-2
 induction therapy. Clones p90 (see X07372) and p98 (see X07373)
 were also obtained. A novel method for cloning cDNAs from mammalian
 expression libraries is based on transient expression of an antigen
 in eukaryotic cells and selection of cells expressing the antigen by
 adhesion to an antibody-coated substrate. The method is useful for
 the isolation and cloning of any protein which can be expressed and
 transported to the cell surface membrane of a eukaryotic cell. It
 has been used to clone genes (see V63442-63) encoding cell surface
 antigens from mammalian lymphocytes (see W80440-55). The isolated
 genes can be expressed in a prokaryotic or eukaryotic host cells to
 produce the encoded protein. The invention also provides high
 efficiency expression vectors (see V63441 and V63444) which allow
 the generation of very large mammalian expression libraries. The
 purified genes and proteins are useful for immunodiagnostic and
 immunotherapeutic applications, including the diagnosis and
 treatment of immune-mediated infections, diseases, and disorders of
 animals, including humans.

SQ Sequence 374 AA;

Query Match 100.0%; Score 57; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYQILTRR 12
 |||||
 Db 241 tsseyqiltrr 252

RESULT 11
 W97833 ID W97833 standard; Protein; 374 AA.
 XX W97833;
 AC


```

XX 07-JUN-1999 (first entry)
DT
XX Human Fc receptor I.
DE
XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
KM cloning.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Misc-difference 2 /note= "encoded by TGG"
FT Misc-difference 23 /note= "encoded by CTC"
FT Misc-difference 44 /note= "encoded by GAC"
FT Misc-difference 45 /note= "encoded by GAC"
FT Misc-difference 60 /note= "encoded by CTC"
FT Misc-difference 60 /note= "encoded by CCC"
FT Misc-difference 77 /note= "encoded by AAT"
FT Misc-difference 85 /note= "encoded by TCC"
FT Misc-difference 99 /note= "encoded by CAA"
FT Misc-difference 103 /note= "encoded by CCC"
FT Misc-difference 141 /note= "encoded by GGC"
FT Misc-difference 159 /note= "encoded by AAC"
FT Misc-difference 171 /note= "encoded by ATG"
FT Misc-difference 176 /note= "encoded by GTC"
FT Misc-difference 256 /note= "encoded by GGC"
FT /note= "encoded by GGC"
PN
XX US5830731-A.
PD
XX 03-NOV-1998.
PF
XX 21-MAY-1997; 97US-0861205.
PR
XX 01-DEC-1992; 92US-0983647.
PR 25-FEB-1988; 88US-0160416.
PR 13-JUL-1989; 89US-0379076.
PR 13-JUL-1990; 90US-0533759.
PR 21-MAY-1997; 97US-0861205.
XX (GEHO ) GEN HOSPITAL CORP.
PA
XX Aruffo A, Seed B;
PI
XX WPI; 1998-609251/51.
DR N-PSDB; X07372.
XX
XX New cloning vector and polylinker - based on existing sequences for
PT efficient cloning and expression of mammalian cDNA(s), especially
PT human lymphocyte antigenic sequences
XX
XX Example 10; Column 53-54; 75bp; English.
XX
XX This is the amino acid sequence of human Fc receptor I (FCRI), as
CC deduced from cDNA clone p90 (see X07372) isolated using a rapid
CC immunoselection cloning method from a cDNA library expressed in COS
CC cells. The cDNA library was constructed from polyA RNA of cells
CC from a single patient undergoing extracorporeal interleukin-2
CC induction therapy. Clones p135 (see V63456) and p98 (see X07373)
CC were also obtained. A novel method for cloning cDNAs from mammalian
CC expression libraries is based on transient expression of an antigen

```

```

CC in eukaryotic cells and selection of cells expressing the antigen by
CC adhesion to an antibody-coated substrate. The method is useful for
CC the isolation and cloning of any protein which can be expressed and
CC transported to the cell surface membrane of a eukaryotic cell. It
CC has been used to clone genes (see V63442-63) encoding cell surface
CC antigens from mammalian lymphocytes (see W80440-55). The isolated
CC genes can be expressed in a prokaryotic or eukaryotic host cells to
CC produce the encoded protein. The invention also provides high
CC efficiency expression vectors (see V63441 and V63444) which allow
CC the generation of very large mammalian expression libraries. The
CC purified genes and proteins are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders of
CC animals, including humans.
XX
XX Sequence 374 AA:
SQ
XX
XX Query Match 100.0%; Score 57; DB 19; Length 374;
XX Best Local Similarity 100.0%; Pred. No. 0.0012;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TSSEYQILTARR 12
XX 241 tsseyqltarr 252
XX
XX RESULT 12
XX ID W97834
XX W97834 standard; Protein; 374 AA.
XX
XX W97834;
XX
XX 07-JUN-1999 (first entry)
DT
XX Human Fc receptor I.
DE
XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
KM cloning.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Misc-difference 2 /note= "encoded by TGG"
FT Misc-difference 23 /note= "encoded by CTC"
FT Misc-difference 44 /note= "encoded by GAC"
FT Misc-difference 45 /note= "encoded by CTC"
FT Misc-difference 60 /note= "encoded by CCC"
FT Misc-difference 77 /note= "encoded by AAT"
FT Misc-difference 85 /note= "encoded by TCC"
FT Misc-difference 99 /note= "encoded by CAA"
FT Misc-difference 103 /note= "encoded by CCC"
FT Misc-difference 141 /note= "encoded by GGC"
FT Misc-difference 159 /note= "encoded by GGC"
FT Misc-difference 171 /note= "encoded by AAC"
FT Misc-difference 176 /note= "encoded by ATG"
FT Misc-difference 256 /note= "encoded by GTC"
FT /note= "encoded by GGC"
PN
XX US5830731-A.

```

XX	03-NOV-1998.
PD	
XX	
PF	21-MAY-1997; 97US-0861205.
XX	
PR	01-DEC-1992; 92US-0983647.
PR	25-FEB-1988; 88US-0160416.
PR	13-JUL-1989; 89US-0379076.
PR	13-JUL-1990; 90US-0553759.
PR	21-MAY-1997; 97US-0861205.
PA	(GENO) GEN HOSPITAL CORP.
XX	
PI	Arufo A., Seed B;
XX	
DR	WPI: 1998-609251/51.
N-PSDB:	X07373.
XX	
PT	New cloning vector and polylinker - based on existing sequences for
PT	efficient cloning and expression of mammalian cDNA(s), especially
PT	human lymphocyte antigenic sequences
XX	
PS	Example 10; Column 53-54; 75pp; English.
CC	This is the amino acid sequence of human Fc receptor I (FcRI), as
CC	deduced from cDNA clone p98/X2 (see X07373) isolated using a rapid
CC	immunoselection cloning method from a cDNA library expressed in COS
CC	cells. The cDNA library was constructed from polyA RNA of cells
CC	from a single patient undergoing extracorporeal interleukin-2
CC	induction therapy. Clones p135 (see V63456) and p90 (see X07372)
CC	were also obtained. A novel method for cloning cDNAs from mammalian
CC	expression libraries is based on transient expression of an antigen
CC	in eukaryotic cells and selection of cells expressing the antigen by
CC	adhesion to an antibody-coated substrate. The method is useful for
CC	the isolation and cloning of any protein which can be expressed and
CC	transported to the cell surface membrane of a eukaryotic cell. It
CC	has been used to clone genes (see V63442-63) encoding cell surface
CC	antigens from mammalian lymphocytes (see W80440-55). The isolated
CC	genes can be expressed in a prokaryotic or eukaryotic host cells to
CC	produce the encoded protein. The invention also provides high
CC	efficiency expression vectors (see V63441 and V63444) which allow
CC	the generation of very large mammalian expression libraries. The
CC	purified genes and proteins are useful for immunodiagnostic and
CC	immunotherapeutic applications, including for diagnosis and
CC	treatment of immune-mediated infections, diseases, and disorders of
CC	animals, including humans.
XX	
SQ	Sequence 374 AA;
XX	
Query Match	100.0%; Score 57; DB 19; Length 374;
Best Local Similarity	100.0%; Pred. NO. 0.0012;
Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 TSSEFOITARR 12
Db	241 tsseyqiltarr 252
RESULT 13	
ID Y96134	Y96134 standard; Protein; 374 AA.
AC XX	Y96134;
DT XX	19-DEC-2000 (first entry)
DE XX	Human macrophage-specific FCRI.
XX	
KM	Macrophage; FCRI; cell surface antigen; human; immunoselection;
KM	panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
KM	immune disorder; infection; asthma; immune-complex disease;
KM	amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.
KW	

XX	Homo sapiens.
XX	
FH	Key
FT	Misc-difference 15 location/Qualifiers
FT	/note= "encoded by GCG"
FT	Misc-difference 38 /note= "encoded by ACC"
FT	Misc-difference 50 /note= "encoded by CGG"
FT	Misc-difference 51 /note= "encoded by ACC"
FT	Misc-difference 55 /note= "encoded by CAC"
FT	Misc-difference 56 /note= "encoded by TTC"
FT	Misc-difference 60 /note= "encoded by CCC"
FT	Misc-difference 64 /note= "encoded by CAG"
FT	Misc-difference 82 /note= "encoded by CAA"
FT	Misc-difference 116 /note= "encoded by GAA"
FT	Misc-difference 117 /note= "encoded by GCA"
FT	Misc-difference 122 /note= "encoded by TTC"
FT	Misc-difference 123 /note= "encoded by ACC"
FT	Misc-difference 126 /note= "encoded by CCG"
FT	Misc-difference 129 /note= "encoded by CAT"
FT	Misc-difference 134 /note= "encoded by AAT"
FT	Misc-difference 136 /note= "encoded by GTT"
FT	Misc-difference 139 /note= "encoded by CCA"
FT	Misc-difference 140 /note= "encoded by AAT"
FT	Misc-difference 213 /note= "encoded by CAA"
FT	Misc-difference 216 /note= "encoded by TTC"
FT	Misc-difference 220 /note= "encoded by CGT"
FT	Misc-difference 268 /note= "encoded by AAT"
FT	Misc-difference 305 /note= "encoded by GTG"
FT	Misc-difference 306 /note= "encoded by AAC"
FT	Misc-difference 332 /note= "encoded by GGT"
PN	US6111093-A.
PD	29-AUG-2000.
PD	28-OCT-1998; 98US-0181612.
PF	01-FEB-1982; 9ZUS-0983647;
PR	25-FEB-1988; 8ZUS-0160416.
PR	13-JUL-1989; 8ZUS-0379076.
PR	23-MAR-1990; 9ZUS-0498809.
PA	(GEHO) GEN HOSPITAL CORP.
PI	Stamenkovic I., Seed B;
XX	

```

DR MPI: 2000-586382/55.
DR N-PSDB; A50592.
XX
XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
XX useful for immunodiagnosis and immunotherapy of immune-mediated
XX infections or disorders, e.g. asthma, immune-complex disease, parasitic
XX diseases
XX
PS Example 10; Column 53-55; 75pp; English.
XX
CC The present sequence is that of a human macrophage specific FCRI,
CC as deduced from cDNA clone p135 (see A50592), which was isolated
CC from a cDNA library expressed in COS cells using a novel method of
CC the invention designed to isolate CSA nucleic acids. The method is
CC based upon transient expression of a CSA in eukaryotic cells and
CC physical selection of cells expressing the antigen by adhesion to
CC (penning on) an antibody-coated substrate such as a culture dish.
CC CSA nucleic acids isolated by the method of the invention, and the
CC proteins they encode, are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders in
CC animals, including humans. These disorders include asthma,
CC immune-complex disease, amyloidosis, parasitic diseases or multiple
CC sclerosis. FCRI is a high affinity receptor for the Fc portion of
CC IgG, normally located on the cell surfaces of macrophages. The
CC ability to interfere with such bonding, or to cause it to occur on
CC surfaces other than macrophages, is useful in therapy. A fusion
CC protein of FCRI and a receptor ligand will be helpful to increase
CC the potencies of antibodies in therapy.
XX
S0 Sequence 374 AA;
XX
Query Match 100.0%; Score 57; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY 1 TSSEYQITARR 12
    |||||
Db 241 tsseyqiltarr 252
RESULT 14
Y96226 Y96226 standard; Protein; 374 AA.
XX ID Y96226 standard; Protein; 374 AA.
XX AC Y96226;
XX Y96226;
XX 11-SEP-2000 (first entry)
XX Human high affinity Fc receptor, FcgammaRI.
XX DE Human high affinity Fc receptor, FcgammaRI.
XX KW Human; high affinity Fc receptor; FcgammaRI; immunoglobulin;
XX infection; immune response; C664; monocyte; macrophage; neutrophil;
XX eosinophil; HIV; IgG; immunosuppressive; antineutritic; cytostatic;
XX antiinflammatory; AIDS; SLE; multiple myeloma; rheumatoid arthritis;
XX systemic lupus erythematosus; tumour.
XX KW
XX Homo sapiens.
XX OS
XX EPI006183-A1.
XX 07-JUN-2000.
XX 03-DEC-1998; 98EP-0122969.
XX PF 03-DEC-1998; 98EP-0122969.
XX PR 03-DEC-1998; 98EP-0122969.
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX MPI: 2000-367968/32.
XX N-PSDB; A27466.
XX

```

Novel Fc receptor lacking transmembrane domains, a signal peptide, and glycosylation, useful for diagnosing and treating immune disorders and cancer -

Disclosure: Page 26-28; 60pp; English.

The present sequence is the human high affinity Fc receptor, Fcgammar1. Fcgammar1 is also known as CD64. Fc receptors play an important role in defending the body against infections. First, pathogens are opsonised by serum immunoglobulins. The resulting complex then binds to cells expressing Fc receptors. Fcgammar1 molecules are expressed by monocytes and macrophages, but expression can also be induced on neutrophils and eosinophils. Upon Fc receptor activation, immune effector pathways are activated, leading to immune response. The present sequence may be modified to produce recombinant versions. The recombinant Fc receptor consist only of the extracellular portion of the receptor and are not glycosylated i.e. they do not have transmembrane domains or signal peptides. The recombinant proteins may be used in immunoassays to determine the immune status of patients with chronic diseases of the immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple myeloma (MM), or rheumatoid arthritis. In addition, pharmaceutical compositions containing recombinant proteins may be used to treat or prevent autoimmune diseases, allergies or tumours, especially AIDS, rheumatoid arthritis or MM.

Sequence 374 AA:

Query Match 100.0%; Score 57; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSEYQILTARR 12
|||||
DB 241 tsseyqiltarr 252

RESULT 15
W86195
W86195 standard; Protein; 377 AA.
W86195;

DT 10-MAY-1999 (first entry)
DE Human Fc receptor I.
XX Fc receptor I; CD36; cell surface antigen; human; cDNA library.
KW Homo sapiens.
OS
FH Key Location/Qualifiers
FT 1..377 /note= "this sequence contains a considerable
FT number of differences from the sequence
FT deduced from the DNA sequence given in
FT the specification"

~US5849898-A.
XX
PD 15-DEC-1998.
XX
PF 07-JUN-1995; 95US-0485447.
XX
PR 01-DEC-1992; 92US-0983647.
PR 25-FEB-1988; 88US-0160416.
PR 13-JUL-1989; 89US-0379076.
PR 23-MAR-1990; 90US-0498809.
PR 13-JUL-1990; 90US-0553759.
PR 07-JUN-1995; 95US-0485447.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX

PI Allen J, Amlot M, Aruffo A, Camerini D, Lauffer L;
 PI Oquendo C, Seed B, Simmons D, Stamenkovic I, Stengelin S;
 XX
 DR WPI: 1999-069813/06.
 DR N-PSDB; V81213.

XX
 PT cDNA encoding human CD40 antigen - useful for cloning cDNA encoding
 PT cell surface antigens, constructing cDNA libraries, expression
 PT vectors for expression in eukaryotic cells or their fragments
 XX

PS Example 10: Column 55-56; 79pp; English.

XX
 CC This polypeptide human Fc receptor 1 (FcRI). FcRI cDNAs (see
 CC V81213) were isolated using a rapid immunoselection cloning
 CC method from a cDNA library expressed in COS cells. The cDNA
 CC library was constructed from polyA RNA of cells from a single
 CC patient undergoing extracorporeal interleukin-2 induction
 CC therapy. DNA sequence analysis revealed that the cDNAs encoded
 CC type I integral membrane proteins with 3 extracellular
 CC immunoglobulin domains. The invention provides a novel method for
 CC cloning cDNAs from mammalian expression libraries. This is based on
 CC transient expression of an antigen in eukaryotic cells and physical
 CC selection of cells expressing the antigen by adhesion to an
 CC antibody-coated substrate. The method is useful for the isolation
 CC and molecular cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell.
 CC CD40_cDNA (see V81198) is specifically claimed.

XX
 SQ Sequence 377 AA;

Query Match 100.0%; Score 57; DB 20; Length 377;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TSSERYQILTARR 12
 |||||||||
 Db 244 tsseyqiltarr 255

Search completed: June 4, 2001, 12:13:30
 Job time: 203 sec

GenCoife version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:30 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-25

Perfect score: 57

Sequence: 1 TSSEYQILFAR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	63.2	91	2	US-08-332-562A-94
2	36	63.2	2004	1	US-08-375-709-15
3	36	63.2	2004	1	US-08-752-929-15
4	36	63.2	2004	1	US-09-090-793-9
5	33	57.9	395	3	US-09-032-372-1
6	33	57.9	1165	1	US-08-240-357-2
7	31	54.4	185	4	US-09-335-409-16
8	31	54.4	451	2	US-08-372-652-2
9	31	54.4	451	5	PCT-US85-16311-2
10	31	54.4	469	3	US-08-372-183-2
11	31	54.4	469	3	US-09-469-721-2
12	31	54.4	469	5	PCT-US95-17023-2
13	31	54.4	472	1	US-08-496-631-2
14	31	54.4	484	2	US-08-372-652-1
15	31	54.4	484	5	PCT-US85-16311-1
16	31	54.4	454	1	US-08-469-486-52
17	31	54.4	454	2	US-08-469-658-52
18	30	52.6	275	4	US-09-036-987A-7
19	30	52.6	326	2	US-08-986-217-3
20	30	52.6	326	2	US-09-154-874-9
21	30	52.6	480	4	US-09-182-859-4
22	30	52.6	480	4	US-09-170-670-5
23	30	52.6	480	4	US-09-193-068-5
24	30	52.6	480	4	US-09-183-412-5
25	30	52.6	483	2	US-08-600-908A-13
26	30	52.6	483	3	US-08-683-838A-13
27	30	52.6	514	1	US-08-720-899-4

28	30	52.6	514	1	US-08-459-610-4	Sequence 4, Appl
29	30	52.6	514	2	US-08-343-804-4	Sequence 4, Appl
30	30	52.6	514	2	US-08-687-399-4	Sequence 4, Appl
31	30	52.6	514	2	US-08-600-908A-4	Sequence 4, Appl
32	30	52.6	514	3	US-08-683-838A-4	Sequence 4, Appl
33	30	52.6	520	1	US-08-468-700-36	Sequence 36, Appl
34	30	52.6	520	1	US-08-645-971-4	Sequence 4, Appl
35	30	52.6	520	2	US-08-468-220-34	Sequence 34, Appl
36	30	52.6	520	2	US-08-468-698-34	Sequence 34, Appl
37	30	52.6	520	2	US-08-704-706A-36	Sequence 36, Appl
38	30	52.6	520	3	US-08-890-383-5	Sequence 5, Appl
39	30	52.6	520	3	US-08-914-679A-5	Sequence 5, Appl
40	30	52.6	520	4	US-08-985-659-37	Sequence 37, Appl
41	30	52.6	520	5	PCT-US94-01553A-34	Sequence 34, Appl
42	30	52.6	520	5	PCT-US95-10426-34	Sequence 34, Appl
43	30	52.6	738	3	US-08-478-208-32	Sequence 32, Appl
44	30	52.6	738	6	5264554-2	Patent No. 5264554
45	30	52.6	904	4	US-09-198-484-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-332-562A-94
Sequence 94, Application US/08332562A
Patent No. 5985599
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/CRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-94
Query Match 63.2% Score/36; DB 2; Length 91;

Best Local Similarity 66.7%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY- 1 TSSEYOILTAR 12
|||||
Db 53 TSSEYHARAER 64

RESULT 2

US-08-375-709-15
; Sequence 15, Application US/08375709
; Patent No. 5683898

GENERAL INFORMATION:

APPLICANT: YAMADA, Kazunaga
APPLICANT: YAMADA, Akiko
APPLICANT: KATO, Seishi
APPLICANT: KONDO, Kiyosi
TITLE OF INVENTION: Gene Coding For Elcosapentaenoic Acid
TITLE OF INVENTION: Synthesizing Enzymes and Process for Production of
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,709
FILING DATE: 20-JAN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/178,251
FILING DATE: 14-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-147945
FILING DATE: 15-MAY-1992

ATTORNEY/AGENT INFORMATION:

NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/150/AAOK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 2004 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-375-709-15

Query Match 63.2%; Score 36; DB 1; Length 2004;
Best Local Similarity 72.7%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSSEYOILTAR 11
|:|:| |
Db 905 TASSEQALRAR 915

RESULT 3

US-08-752-929-15
; Sequence 15, Application US/08752929
; Patent No. 5798259

GENERAL INFORMATION:

APPLICANT: YAMADA, Kazunaga
APPLICANT: YAMADA, Akiko
APPLICANT: KATO, Seishi
APPLICANT: KONDO, Kiyosi
TITLE OF INVENTION: Gene Coding for Elcosapentaenoic Acid Synthesizing
TITLE OF INVENTION: Enzymes and Process for Production of Elcosapentaenoic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,929
FILING DATE: 20-NOV-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/375,709
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,251
FILING DATE: 14-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-147945
FILING DATE: 15-MAY-1992

ATTORNEY/AGENT INFORMATION:

NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/150/AAOK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 2004 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-752-929-15

Query Match 63.2%; Score 36; DB 1; Length 2004;
Best Local Similarity 72.7%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSSEYOILTAR 11
|:|:| |
Db 905 TASSEQALRAR 915

RESULT 4

US-09-090-793-9
; Sequence 9, Application US/09090793
; Patent No. 6140486

GENERAL INFORMATION:

APPLICANT: Calgene, LLC
TITLE OF INVENTION: Production of polyunsaturated fatty acids by expression
TITLE OF INVENTION: of polyketide-like synthesis genes in plants
FILE REFERENCE: CGNE.131.0105
CURRENT APPLICATION NUMBER: US/09/090,793
CURRENT FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,650
EARLIER FILING DATE: 1997-06-04

NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 9
LENGTH: 2004
TYPE: PRT
ORGANISM: Shewanella putrefaciens
US-09-090-793-9

Query Match 63.2%; Score 36; DB 4; Length 2004;
Best Local Similarity 72.7%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSSEYOILTAR 11
DB 905 TASEOQALTR 915

RESULT 5
US-09-032-372-1
Sequence 1, Application US/09032372
Patent No. 6008337
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Yue, Henry
TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,372
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0478 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAB01
CLONE: 78191
US-09-032-372-1

Query Match 57.9%; Score 33; DB 3; Length 395;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSSEYOILTA 10
DB 204 TNSDMOILTA 213

RESULT 6
US-08-240-357-2
Sequence 2, Application US/08240357
Patent No. 5578481
GENERAL INFORMATION:
APPLICANT: Ishikawa, Yoshihiro
TITLE OF INVENTION: Cloning and Characterization of a
TITLE OF INVENTION: Cardiac Adenylyl Cyclase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470-8426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,357
FILING DATE: 10-MAY-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31,705-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1165 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-240-357-2

Query Match 57.9%; Score 33; DB 1; Length 1165;
Best Local Similarity 50.0%; Pred. No. 14e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSSEYOILTARR 12
DB 1126 TTLDYOVLAAR 1137

RESULT 7
US-09-335-409-16
Sequence 16, Application US/09335409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHIIONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 16.
LENGTH: 185
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-16

Query Match 54.4%; Score 31; DB 4; Length 185;
Best Local Similarity 63.6%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 TSSEYOILTA 12
DB 158 ASDQOILTA 168

RESULT 8
US-08-372-652-2
Sequence 2, Application US/08372652;
Patent No. 5932699.

GENERAL INFORMATION:
APPLICANT: Moore, David
APPLICANT: Seol, Wongi
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street, Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/372,652
FILING DATE: 13-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/246001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-372-652-2

Query Match 54.4%; Score 31; DB 2; Length 451;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTA 10
DB 357 TOEYALTA 366

RESULT 9
PCT-US95-16311-2
Sequence 2, Application PC/TUS9516311
GENERAL INFORMATION:

APPLICANT: Moore, David
APPLICANT: Seol, Wongi
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street, Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16311
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/372,652
FILING DATE: 13-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/246001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-16311-2

Query Match 54.4%; Score 31; DB 5; Length 451;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTA 10
DB 357 TOEYALTA 366

RESULT 10
US-08-372-183-2
Sequence 2, Application US/08372183
Patent No. 6005086

GENERAL INFORMATION:
APPLICANT: Evans, Ronald M.
APPLICANT: Portan, Barry M.
APPLICANT: Weinberger, Cary A.
TITLE OF INVENTION: METHOD FOR MODULATING PROCESSES MEDIATED
BY FARNESOID ACTIVATED RECEPTORS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/372,183
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9844
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-372-183-2

Query Match 54.4%; Score 31; DB 3; Length 469;
Best Local Similarity 60.0%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOLTA 10
| 11:111
Db 375 TOEYALLTA 384

RESULT 11
US-09-469-721-2
Sequence 2, Application US/09469721
Patent No. 6184353
GENERAL INFORMATION:
APPLICANT: Evans, Ronald M.
APPLICANT: Forman, Barry M.
APPLICANT: Weinberger, Cary A.
TITLE OF INVENTION: METHOD FOR MODULATING PROCESSES MEDIATED
TITLE OF INVENTION: BY FARNESOID ACTIVATED RECEPTORS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/469,721
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/372,183
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9844
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-469-721-2

Query Match 54.4%; Score 31; DB 4; Length 469;
Best Local Similarity 60.0%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOLTA 10
| 11:111
Db 375 TOEYALLTA 384

RESULT 12
PCT-US95-17023-2
Sequence 2, Application PC/TUS9517023

GENERAL INFORMATION:
APPLICANT: Evans, Ronald M.
APPLICANT: Forman, Barry M.
APPLICANT: Weinberger, Cary A.
TITLE OF INVENTION: METHOD FOR MODULATING PROCESSES MEDIATED
TITLE OF INVENTION: BY FARNESOID ACTIVATED RECEPTORS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17023
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9844
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-17023-2

Query Match 54.4%; Score 31; DB 5; Length 469;
Best Local Similarity 60.0%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOLTA 10
| 11:111
Db 375 TOEYALLTA 384

RESULT 13
US-08-496-631-2
Sequence 2, Application US/08496631

GENERAL INFORMATION:
APPLICANT: Bowman, Michael
APPLICANT: Reiter, Stephen E.
TITLE OF INVENTION: STEROID RECEPTOR RRI
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,631
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-496-631-2

Query Match 54.4%; Score 31; DB 1; Length 472;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSSEYQILTA 10
DB 378 TOEYVALTA 387

RESULT 14

US-08-372-652-1
Sequence 1, Application US/08372652
Patent No. 5932699
GENERAL INFORMATION:
APPLICANT: Moore, David
APPLICANT: Seol, Mongi
APPLICANT: Choi, Hwang-Sik
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street, Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/372,652
FILING DATE: 13-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/246001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154

INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-372-652-1

Query Match 54.4%; Score 31; DB 2; Length 484;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSSEYQILTA 10
DB 390 TOEYVALTA 399

RESULT 15

PCT-US95-16311-1
Sequence 1, Application PC/TUS9516311
GENERAL INFORMATION:
APPLICANT: Moore, David
APPLICANT: Seol, Mongi
APPLICANT: Choi, Hwang-Sik
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street, Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16311
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/372,652
FILING DATE: 13-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/246001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-16311-1

Query Match 54.4%; Score 31; DB 5; Length 484;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSSEYQILTA 10
DB 390 TOEYVALTA 399

Tue Jun 5 07:09:04 2001

Search completed: June 4, 2001, 12:14:31
Job time: 224 sec

us-09-284-107-25.ra1

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:42 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-25

Perfect score: 57
Sequence: 1 TSSEYQILFARR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 6872935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :

PIR-67:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	57	100.0	344	2	A41357	Fc gamma (Igg) rec
2	57	100.0	374	1	A39878	Fc gamma (Igg) rec
3	36	63.2	336	2	I48471	Fc gamma (Igg) rec
4	36	63.2	404	2	A46480	Fc gamma (Igg) rec
5	36	63.2	409	2	H83382	probable porin PA2
6	36	63.2	2004	2	T30185	hypothetical prote
7	35	61.4	197	2	T15725	hypothetical prote
8	35	61.4	433	2	S46668	MTM1 protein - yea
9	35	61.4	880	2	S60137	beta-N-acetylhexos
10	34	59.6	277	2	T08554	hypothetical prote
11	34	59.6	475	2	F70972	probable amidase -
12	34	59.6	863	2	T49709	related to glucan
13	33	57.9	238	2	C83240	probable transcrip
14	33	57.9	258	2	T36303	probable indolegly
15	33	57.9	271	2	T50620	hypothetical prote
16	33	57.9	278	2	A49505	lectin light chain
17	33	57.9	288	2	A48452	galactose-inhibita
18	33	57.9	305	2	S75743	methionyl aminopep
19	33	57.9	353	2	A54969	beta-5-GTP-binding
20	33	57.9	413	2	T21644	hypothetical prote
21	33	57.9	484	1	NNBY2	anthranilate synth
22	33	57.9	484	2	G83456	xanthine dehydroge
23	33	57.9	620	2	A61086	dopamine beta-mono
24	33	57.9	696	2	G71829	probable outer mem
25	33	57.9	752	2	S30084	anthranilate synth
26	33	57.9	762	1	NNNC2	anthranilate synth
27	33	57.9	768	2	S04518	anthranilate synth
28	33	57.9	768	2	S07305	trpC protein - Eme
29	33	57.9	770	2	S00643	anthranilate synth

30	33	57.9	770	2	S11161	anthranilate synth
31	33	57.9	770	2	T07692	hypothetical prote
32	33	57.9	887	2	S73768	Me277 homolog F11
33	33	57.9	1165	2	A46180	adenyl cyclase t
34	32	56.1	131	2	G70388	hypothetical prote
35	32	56.1	156	2	A69966	hypothetical prote
36	32	56.1	182	2	T21316	hypothetical prote
37	32	56.1	212	2	T30700	hypothetical prote
38	32	56.1	429	2	T50362	cdcl-1-like protein
39	32	56.1	433	2	D75480	UDP-N-acetylmuramo
40	32	56.1	520	1	S02505	nitrogenase (Ec 1.
41	32	56.1	570	2	F70844	probable fusion pr
42	32	56.1	597	2	C83129	hypothetical prote
43	32	56.1	623	1	VCNMBE	spike glycoprotein
44	32	56.1	731	2	T08855	nephrocytin - hum
45	32	56.1	969	2	F71418	hypothetical prote

ALIGNMENTS

```

RESULT 1
Fc gamma (Igg) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C>Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000
C:Accession: A41357; S03019
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A:Title: Isolation and expression of functional high-affinity Fc receptor complementa
A:Reference number: A41357; M0ID:89100284
A:Accession: A41357
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: GB:X14355; GB:M21090; NID:G31333; PIDN:CAA32536.1; PID:G31334
R:Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (
A:Reference number: S03018; M0ID:89098339
A:Accession: S03019
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: EMBL:X14355; NID:G31333; PIDN:CAA32536.1; PID:G31334
A:Note: the authors translated the codon ACG for residue 25 as Ala
C:Superfamily: Fc gamma receptor I; Immunoglobulin homology
C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
F:117-170/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 57; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSSEYQILFARR 12
DB 241 TSSEYQILFARR 252

RESULT - 2
Fc gamma (Igg) receptor I-A (high affinity) precursor - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C>Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: A39878; I70304; B41357; S03018; I57525
R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
J. Biol. Chem. 266, 13449-13455, 1991
A:Title: Gene organization of the human high affinity receptor for Igg, Fc gammaRI (CD
A:Reference number: A39878; M0ID:91302383
A:Accession: A39878
A:Molecule type: DNA
A:Residues: 1-374 <VAN>

```

A:Cross-references: GB:M63830; GB:M63835; NID:9180277; PIDN:AAA5678.1; PID:9180279
 R:Forges, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: I55577; MUID:93055454
 A:Accession: 170304
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:L03418; NID:9184840; PIDN:AAA36049.1; PID:9292169
 R:Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: A41357; MUID:89100284
 A:Accession: B41357
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', 26-337, 'T', 339-374 <AL1>
 A:Cross-references: GB:X14356; GB:M21091; NID:931331; PIDN:CAA32537.1; PID:931332
 R:Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR)
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', 26-337, 'T', 339-374 <AL2>
 A:Cross-references: EMBL:X14356; NID:931331; PIDN:CAA32537.1; PID:931332
 R:Perez, C.; Wietzerbin, J.; Benech, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A:Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
 1sm.
 A:Reference number: I57525; MUID:93204964
 A:Accession: I57525
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RE2>
 A:Cross-references: GB:SS7204; NID:9298692; PIDN:AAD13887.1; PID:94261587
 C:Genetics:
 A:Gene: GDB:FCGR1A; CD64
 A:Cross-references: GDB:135911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane P
 F:1-13/Domain: signal sequence #status: predicted <SIG>
 F:16-292/Domain: extracellular #status: predicted <EXT>
 F:117-170/Domain: immunoglobulin homology <IMM2>
 F:293-313/Domain: transmembrane #status: predicted <IMM>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent) #status: predicted

Query Match 100.0%; Score 57; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0063;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy- 1 TSSEYQILTARR 12
 |||||
 DB 241 TSSEYQILTARR 252

RESULT 3
 148471
 Fc gamma (IgG) receptor high affinity - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
 C:Accession: I48471
 R:Pirius, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarth, P.M.; Wicker, L.S.; Gaff
 Science 260, 695-698, 1993
 A:Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for Ig
 A:Reference number: I48471; MUID:93242399
 A:Accession: I48471
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-336 <RES>
 A:Cross-references: EMBL:X70980; NID:9311748; PIDN:CAA50311.1; PID:9311749

C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: immunoglobulin receptor
 F:128-180/Domain: immunoglobulin homology <IMM>

Query Match 63.2%; Score 36; DB 2; Length 336;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 TSSEYQILTARR 12
 |||||
 DB 255 TSSEYHARAER 266

RESULT 4
 A46480
 Fc gamma (IgG) receptor high affinity - mouse
 N:Alternate names: high affinity IgG receptor
 C:Species: Mus musculus (house mouse)
 C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
 C:Accession: A46480; A43511
 R:Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.
 J. Immunol. 148, 1570-1575, 1992
 A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI a
 A:Reference number: A46480; MUID:92166399
 A:Accession: A46480
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-404 <OSM>
 A:Note: sequence extracted from NCBI Backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212,
 R:Seares, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.
 J. Immunol. 144, 371-378, 1990
 A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for
 A:Reference number: A43511; MUID:90111035
 A:Accession: A43511
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-404 <SEA>
 A:Cross-references: GB:M41314; NID:9200752; PIDN:AAA40056.1; PID:9200753
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: immunoglobulin receptor; transmembrane protein
 F:127-179/Domain: immunoglobulin homology <IMM>

Query Match 63.2%; Score 36; DB 2; Length 404;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 TSSEYQILTARR 12
 |||||
 DB 250 TSSEYHARAER 261

RESULT 5
 H83382
 Probable porin PA2113 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: H83382
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337
 A:Accession: H83382
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-409 <STO>
 A:Cross-references: GB:AE004638; GB:AE004091; NID:99948116; PIDN:AAG05501.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2113

Query Match 63.2%; Score 36; DB 2; Length 409;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 TSSEYOILTRAR 12
 |||:|:|:|:|
 DB 181 SSDDYOFVFSNR 192

RESULT 6
 T30185
 hypothetical protein 7 - Shewanella sp.
 C:Species: Shewanella sp.
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C:Accession: T30185
 R:Takeyama, H.; Takeda, D.; Yazawa, K.; Yamada, A.; Matsunaga, T.
 Microbiology 143, 2725-2731, 1997
 A:Title: Expression of the elcosapentaenoic acid synthesis gene cluster from Shewanella
 A:Reference number: Z20764; MUID:97419510
 A:Accession: T30185
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2004 <TAK>
 A:Cross-references: EMBL:U73935; NID:g2529413; PID:g2529420; PIDN:AA81125.1
 A:Experimental source: strain SCRC-2738

Query Match 63.2%; Score 36; DB 2; Length 2004;
 Best Local Similarity 72.7%; Pred. No. 80;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TSSEYOILTRAR 11
 |||:|:|:|
 DB 905 TASEOQALTRAR 915

RESULT 7
 T15725
 hypothetical protein C31H1.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 04-Mar-2000
 C:Accession: T15725
 R:Le, J.
 submitted to the EMBL Data Library, December 1995
 A:Description: The sequence of C. elegans cosmid C31H1.
 A:Reference number: Z18395
 A:Accession: T15725
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-197 <LET>
 A:Cross-references: EMBL:U42848; NID:g1125817; PID:g1125818; PIDN:AAA83606.1; CESP:C31H1
 C:Genetics:
 A:Gene: CESP:C31H1.2
 A:Introns: 21/1: 98/3; 130/3; 143/1; 178/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C31H1.2

Query Match 61.4%; Score 35; DB 2; Length 197;
 Best Local Similarity 58.3%; Pred. No. 11;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 TSSEYOILTRAR 12
 |||:|:|:|:|
 DB 23 TSSEIOLLSMKR 34

RESULT 8
 S4668
 MTH1 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D9954.12; protein YDR277c
 C:Species: Saccharomyces cerevisiae

C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 21-Jul-2000
 C:Accession: S4668; S70133
 R:Hubbard, E.J.A.; Jiang, R.; Carlson, M.
 Mol. Cell. Biol. 14, 1972-1978, 1994
 A:Title: Dosage-dependent modulation of glucose repression by MSN3 (STD1) in Saccharo
 A:Reference number: S4668; MUID:94158870
 A:Accession: S4668
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-433 <HUB>
 A:Cross-references: EMBL:L21933; NID:g416159; PIDN:AAA18536.1; PID:g416160
 R:Le, T.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of S. cerevisiae cosmid 9954.
 A:Reference number: S70124
 A:Accession: S70133
 A:Molecule type: DNA
 A:Residues: 1-433 <LET>
 A:Cross-references: EMBL:U51030; NID:g1332633; PID:g1230646; MIPS:YDR277c
 C:Genetics:
 A:Gene: SGD:MTH1
 A:Cross-references: SGD:S0002685; MIPS:YDR277c
 A:Map position: 4R

Query Match 61.4%; Score 35; DB 2; Length 433;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 6; Conservative 4; Mismatches 2; Indels ~0; Gaps 0;

OY 1 TSSEYOILTRAR 12
 |||:|:|:|:|
 DB 273 TSATYVATARK 284

RESULT 9
 S60137
 beta-N-acetylhexosaminidase (EC 3.2.1.52) lyTD precursor - Bacillus subtilis
 N:Alternate names: major autolysin lyTD; N-acetylglucosaminidase
 C:Species: Bacillus subtilis
 C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 15-Oct-1999
 C:Accession: S60137; S49390; F69654
 R:Margot, P.; Maue, C.; Karamata, D.
 submitted to the EMBL Data Library, October 1993
 A:Description: Cloning, sequencing and disruption of the N-acetylglucosaminidase gene
 A:Reference number: S60137
 A:Accession: S60137
 A:Molecule type: DNA
 A:Residues: 1-880 <MAR>
 A:Cross-references: EMBL:U02562; NID:g476091; PIDN:AA67857.1; PID:g476093
 A:Experimental source: strain 168
 R:Margot, P.; Maue, C.; Karamata, D.
 Mol. Microbiol. 12, 535-545, 1994
 A:Title: The gene of the N-acetylglucosaminidase, a Bacillus subtilis 168 cell wall h
 A:Reference number: S49389; MUID:95020588
 A:Accession: S49390
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 81-132;161-212;439-473;479-513 <MAN>
 A:Cross-references: EMBL:U02562
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emerson, P.T.; Ertian, K.D.; Errington, J.; Fabret, C.; Ferrari
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 Koeltter, P.; Koningsstein, G.; Krogh, S.; Kunita, M.; Kurita, K.; Lapidos, A.; Lardino
 A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 Y., M.; Ogawa, K.; Ogiwara, A.; Ouega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portee
 Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, W.; Sadate, Y.; Sato, T.; Scanl
 A:Authors: Schleich, S.; Schreier, R.; Scrofano, P.; Sekiguchi, J.; Sekowska, A.; Se
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MID:98044033
A:Accession: F69654
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-880 <RUN>
A:Cross-references: GB:299122; GB:AL009126; NID:g2636029; PIDN:CAB15595.1; PID:el184484;
A:Experimental source: strain 168
C:Genetics:
A:Gene: lytD
C:Keywords: glycosidase; hydrolase
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-880/Product: beta-N-acetylhexosaminidase #status predicted <MAT>

Query Match 61.4%; Score 35; DB 2; Length 880;
Best Local Similarity 63.6%; Pred. No. 54;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 11
||:|||||:
Db 353 TSAAYOITAK 363

RESULT 10
T08554
hypothetical protein F27B13.170 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C:Accession: T08554
R:Bevan, M.; Zimmermann, W.; Gruenisehn, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16442
A:Accession: T08554
A:Molecule type: DNA
A:Residues: 1-277 <BEV>
A:Cross-references: EMBL:AL050352; GSPDB:GN00062; ATSP:F27B13.170
A:Experimental source: cultivar Columbia; BAC clone F27B13
C:Genetics:
A:Gene: ATSP:F27B13.170
A:Map position: 4
A:introns: 27/2; 82/3; 168/3

Query Match 59.6%; Score 34; DB 2; Length 277;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
||:|||||:
Db 257 TSTNYKILNASR 268

RESULT 11
F70972
probable amidase - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70972
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: F70972
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-475 <COL>
A:Cross-references: GB:AL009198; GB:AL123456; NID:g3242262; PIDN:CAA15760.1; PID:g266165
A:Experimental source: strain H37RV
C:Genetics:

A:Gene: amid
C:Superfamily: indoleacetamide hydrolase

Query Match 59.6%; Score 34; DB 2; Length 475;
Best Local Similarity 77.8%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 EYOILTARR 12
||| ||| ||
Db 360 EYOITERR 368

RESULT 12
T49709
related to glucan 1,3-beta-glucosidase precursor [imported] - *Neurospora crassa*
N:Alternate names: protein B23L21.220
C:Species: *Neurospora crassa*
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000
C:Accession: T49709
R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49709
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-863 <SCH>
A:Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.220
A:Experimental source: BAC clone B23L21; strain OR7A
C:Genetics:
A:Gene: NCSP:B23L21.220
A:Map position: 6

Query Match 59.6%; Score 34; DB 2; Length 863;
Best Local Similarity 50.0%; Pred. No. 85;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
||| ||| |||:
Db 228 SDPEYOITARR 239

RESULT 13
C83240
probable transcription regulator PA3249 [imported] - *Pseudomonas aeruginosa* (strain P
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83240
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: C83240
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <STO>
A:Cross-references: GB:AE004747; GB:AE004091; NID:g9949362; PIDN:AA006637.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3249

Query Match 57.9%; Score 33; DB 2; Length 238;
Best Local Similarity 72.7%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSEYOILTARR 12
||| ||| |||
Db 123 SSVOIRARR 133

RESULT 14

probable indoleglycerol phosphate synthase - Streptomyces coelicolor
T36303
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C:Accession: T36303
R:Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21604
A:Accession: T36303
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-258 <SAD>
A:Cross-references: EMBL:AL035654; PIDN:CAB38582.1; GSPDB:GN00070; SCOEDB:SCE8.04C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: trpC2; SCOEDB:SCE8.04C
C:Superfamily: indole-3-glycerol-phosphate synthase; trpC homology

Query Match

Best Local Similarity 57.9%; Score 33; DB 2; Length 258;
Pred. No. 38;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 EYQILFAR 11
|||||
DB-- 123 EYQILFAR 130

RESULT 15

T50620
hypothetical protein DKFZp762M186.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50620
R:Blöcker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Well, B.; Wiemann, S.
submitted to the Protein Sequence Database, June 2000
A:Reference number: Z25143
A:Accession: T50620
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-271 <AAA>
A:Cross-references: EMBL:AL359590
A:Experimental source: adult melanoma (Mewo cell line); clone DKFZp762M186
C:Genetics:
A:Note: DKFZp762M186.1

Query Match 57.9%; Score 33; DB 2; Length 271;

Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSSEYOILTA 10
|:|||||
DB 117 TRAEYOILTA 126

Search completed: June 4, 2001, 12:15:43
Job time: 276 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:05 ; Search time 37.56 Seconds
(without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-25
Sequence: 1 TSSEQILFARR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	374	1	FCG1_HUMAN
2	56	63.2	404	1	FCG1_MOUSE
3	35	61.4	433	1	MTM1_YEAST
4	35	61.4	880	1	LYPD_BACSU
5	34	59.6	475	1	AM14_MYCTU
6	33	57.9	258	1	TRC2_STRCO
7	33	57.9	288	1	GIL2_ENTHI
8	33	57.9	293	1	PAR3_CAVCR
9	33	57.9	305	1	AMP3_SYNY3
10	33	57.9	353	1	GBB5_HUMAN
11	33	57.9	333	1	GBB5_MOUSE
12	33	57.9	413	1	YV4Q_CAEEL
13	33	57.9	484	1	TRPG_YEAST
14	33	57.9	620	1	DOPO_RAT
15	33	57.9	682	1	CJBA_BACUH
16	33	57.9	752	1	TRPG_PENCH
17	33	57.9	762	1	TRPG_NEOCR
18	33	57.9	768	1	TRPG_ASPAW
19	33	57.9	768	1	TRPG_EMENT
20	33	57.9	770	1	TRPG_ASPNG
21	33	57.9	826	1	CRAA_BACUH
22	33	57.9	887	1	Y277_MYCPN
23	33	57.9	1185	1	CYAB_CANFA
24	33	57.9	1185	1	Y0Y1_BACSU
25	32	56.1	456	1	FTSA_PORGI
26	32	56.1	490	1	FA10_RABIT
27	32	56.1	520	1	NIEK_KLEPN
28	32	56.1	623	1	VG16_BREV
29	32	56.1	777	1	BISC_ECOLI
30	31	54.4	284	1	AMPE_ECOLI
31	31	54.4	301	1	Y186_MYCPN
32	31	54.4	322	1	HCR_ECOLI
33	31	54.4	349	1	US30_HCMVA

34	31	54.4	403	1	T230_CAEEL	009474 caenorhabd
35	31	54.4	437	1	FLHF_PSEPU	052256 pseudomonas
36	31	54.4	444	1	STDL_YEAST	002724 saccharomyc
37	31	54.4	464	1	HSLD_BACHD	09ka27 bacillus ha
38	31	54.4	477	1	TRPG_HAEIN	046451 haemophilus
39	31	54.4	480	1	HEP2_RABIT	047776 oryctolagus
40	31	54.4	567	1	CH13_CANAL	040954 candida alb
41	31	54.4	680	1	SYG_BOMMO	004451 bombyx mori
42	31	54.4	699	1	YA76_SCHPO	009761 schizosacch
43	31	54.4	707	1	PNP_BUCAT	057454 buchnera ap
44	31	54.4	741	1	HOX1_HAIRO	028468 halocynthia
45	31	54.4	794	1	HMMR_MOUSE	000547 mus musculu

ALIGNMENTS

RESULT 1
FCG1_HUMAN STANDARD; PRT; 374 AA.
ID FCG1_HUMAN
AC P12314; P12315;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, last sequence update)
DT 01-OCT-2000 (Rel. 40, last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA
DE RI) (FCRI) (IGG FC RECEPTOR I) (CD64).
GN FCGRIA OR FCGRI OR FCGI OR IGFRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89098339; PubMed=2974947;
RA Allen J.M., Seed B.;
RT "Nucleotide sequence of three cDNAs for the human high affinity Fc
RT receptor (FCRI).";
RL Nucleic Acids Res. 16:11824-11824(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89100284; PubMed=2911749;
RA Allen J.M., Seed B.;
RT "Isolation and expression of functional high-affinity Fc receptor
RT complementary DNAs.";
RL Science 243:378-381(1989).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH
CC -1- AFFINITY RECEPTOR.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD64 entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prow/cd/cd64.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X14356; CAA32537.1;
CC EMBL: X14355; CAA32536.1;
CC PIR: S03018; S03018.
CC PIR: S03019; S03019.
CC PIR: A41357; A41357.
CC PIR: B41357; B41357.
CC HSSP: P12319; 1ALT.
CC MIM: 146760;
CC InterPro: IPR003006;
CC Pfam: PF00047; 1g; 3.

KW 196-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KM Immunoglobulin domain; Alternative splicing; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 2 374
 FT DOMAIN 22 292
 FT TRANSMEM 293 313
 FT DOMAIN 314 374
 FT DOMAIN 36 92
 FT DOMAIN 117 175
 FT DOMAIN 205 267
 FT DISULFID 43 85
 FT DISULFID 124 168
 FT DISULFID 212 260
 FT CARBOHYD 59 78
 FT CARBOHYD 152 152
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT VARSPLIC 333 374
 FT VARIANT 26 26
 FT VARIANT 59 59
 FT VARIANT 59 59
 FT CONFLICT 25 25
 FT SEQUENCE 374 AA; 42605 MW; 2C2A8A103EECF1E6 CRC64;

Query Match 100.0%; Score 57; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSSEYOILTARR 12
 DB 241 TSSEYOILTARR 252

RESULT 2
 FCGL_MOUSE STANDARD; PRT; 404 AA.
 AC P26151;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I)
 GN FCGR1 OR FCGL
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90111035; PubMed-2136886;
 RA Seads D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RT "Molecular cloning and expression of the mouse high affinity Fc receptor for IgG";
 RL J. Immunol. 144:371-378(1990).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92166399; PubMed-1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RT "Structure and mapping of the gene encoding mouse high affinity Fc gamma RI and chromosomal location of the human Fc gamma RI gene";
 RL J. Immunol. 148:1570-1575(1992).
 RN (3)
 RP FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN GAMMA. HIGH AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M31314; AAA40056.1;
 DR PIR: A43511; A43511.
 DR PIR: A46480; A46480.
 DR HSSP: P12319; IALT.
 DR MGD: MGI:95498; FCgr1.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; Ig_3.
 KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KM Immunoglobulin domain.
 FT SIGNAL 1 24
 FT CHAIN 25 404
 FT DOMAIN 25 297
 FT TRANSMEM 298 320
 FT DOMAIN 321 404
 FT DOMAIN 46 66
 FT DOMAIN 67 124
 FT DOMAIN 154 216
 FT CARBOHYD 28 28
 FT CARBOHYD 48 48
 FT CARBOHYD 69 69
 FT CARBOHYD 168 168
 FT CARBOHYD 249 249
 FT SEQUENCE 404 AA; 44887 MW; 1CAFD0033842767E7 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 404;
 Best Local Similarity 66.7%; Pred. No. 5.5;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TSSEYOILTARR 12
 DB 250 TSSEYHIAARR 261

RESULT 3
 MTHL_YEAST STANDARD; PRT; 433 AA.
 AC P35198;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MTHL PROTEIN.
 GN MTHL OR YDR277C OR D9954.12.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX STRAIN-S288C;
 RX MEDLINE-94158870; PubMed-8114728;
 RA Hubbard E.J.A., Jiang R., Carlson M.;
 RT "Dosage-dependent modulation of glucose repression by MSN3 (STO1) in
 RL Saccharomyces cerevisiae";
 RL Mol. Cell. Biol. 14:1972-1978(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX STRAIN-S288C / AB972;
 RC Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Langston V., Latreille P., Le T.,

RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifkin L., Riles L., Tatch A., Trevasakis E., Vignati D.,
 RA Wilcox L., Wolfdman P., Vaudin M., Wilson R., Waterston R.,
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: STRONG, TO STD/MSN3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L21933; AAA18536.1; -
 DR EMBL: U51030; AAB64457.1; -
 DR PIR: S4668; S4668.
 DR SGD: S0002685; MTH1.
 SQ SEQUENCE 433 AA; 49060 MW; 3A23EF4742632761 CRC64;
 Query Match 61.4%; Score 35; DB 1; Length 433;
 Best Local Similarity 50.0%; Pred. No. 9.7;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TSSEYQILTRR 12
 DB 273 TSAKYTVATARK 284
 RESULT 4
 ID LYTID_BACSU STANDARD; PRT; 880 AA.
 AC P39848;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BETA-N-ACETYLGLUCOSAMINIDASE PRECURSOR (EC 3.2.1.-).
 GN LYTID OR CWLG.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95020588; PubMed-7934877;
 RA Margot P., Maueel C., Karamata D.;
 RT "The gene of the N-acetylglucosaminidase, a Bacillus subtilis 168
 RT cell wall hydrolase not involved in vegetative cell autolysis.";
 RL M01. Microbiol. 12:535-545(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AC327;
 RX MEDLINE=96036197; PubMed-7581999;
 RA Rashid M.H., Mori M., Sekiguchi J.;
 RT "Glucosaminidase of Bacillus subtilis: cloning, regulation, primary
 RT structure and biochemical characterization.";
 RL Microbiology 141:2391-2404(1995).
 CC -1- FUNCTION: CELL WALL HYDROLASE NOT INVOLVED IN CELL AUTOLYSIS. IT
 CC HYDROLYZES THE BETA-1,4 GLYCAN BOND BETWEEN THE N-
 CC ACETYLGLUCOSAMINYL AND THE N-ACETYLGLUCOSAMINOYL RESIDUES IN THE GLYCAN
 CC CHAIN.
 CC SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL: U02562; AAA67857.1; -
 DR EMBL: D45046; BAA08089.1; -
 DR EMBL: Z99122; CAB15595.1; -
 DR Subtilisin; BG10455; lytd.
 DR InterPro: IPR002901; -
 DR Pfam: PF01832; Amidase_4; 1.
 KW Cell wall; Hydrolase; Signal; Repeat.
 FT SIGNAL 1 27
 FT CHAIN 28 880
 FT DOMAIN 72 75
 FT DOMAIN 337 340
 FT DOMAIN 568 571
 FT REPEAT 439 473
 FT REPEAT 479 513
 FT REPEAT 513 513
 SQ SEQUENCE 880 AA; 95553 MW; 2A912A478FCFC1D1 CRC64;
 Query Match 61.4%; Score 35; DB 1; Length 880;
 Best Local Similarity 63.6%; Pred. No. 21;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TSSEYQILTRR 11
 DB 353 TSAAYQITTK 363
 RESULT 5
 ID AM14_MYCTU STANDARD; PRT; 475 AA.
 AC O50404;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE AMIDASE AMID (EC 3.5.1.4).
 GN AMID OR RV3375 OR MTW004.33.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Sultison J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -1- CATALYTIC ACTIVITY: A MONOCARBOXYLIC ACID AMIDE + H(2)O -
 CC -1- A MONOCARBOXYLATE + NH(3).
 CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AL009198; CA15760.1; -
 DR TubercuList; RV3375; -
 DR InterPro: IPR000120; -
 DR Pfam: PF01425; Amidase; 1.
 DR PROSITE: PS00571; AMIDASES; FALSE_NEG.
 KW Hypothetical protein; Hydrolase.

SQ SEQUENCE 475 AA; 50645 MW; 8A74C138EE3C6745 CRC64;
 Query Match 59.6%; Score 34; DB 1; Length 475;
 Best Local Similarity 77.8%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 EYOILTFARR 12
 ||| ||| ||
 Db 360 EYOILTFARR 368
 RESULT 6
 TRC2_STRCO STANDARD; PRT; 258 AA.
 AC 0924X0;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE 2 (EC 4.1.1.48) (IGPS).
 GN TRPC2 OR SCE8.04C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_Taxid=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3(2);
 RA Saunders D., Harris D., Bentley S.D., Parkhill J., Barrell B.G.,
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE FUNCTION OF THE SECOND TRP OPERON IN S. COELICOLOR IS
 TO PRODUCE TRP FOR THE BIOSYNTHESIS OF CALCIUM-DEPENDENT
 ANTIBIOTIC (CDA).
 CC -1- CATALYTIC ACTIVITY: 1-(2-CARBOXYPHENYLAMINO)-1-DEOXY-D-RIBULOSE
 5-PHOSPHATE -> 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE + CO(2) + H(2)O.
 CC -1- PATHWAY: FOURTH STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
 CC -1- SIMILARITY: BELONGS TO THE TRPC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL: AL035654; CAB38582.1;
 DR HSSP: P00909; IPII
 DR InterPro: IPR001468;
 DR Pfam: PF00218; IGPS; 1;
 DR PROSITE: PS00614; IGPS; 1;
 KM Tryptophan biosynthesis; Lyase; Decarboxylase;
 KN Antibiotic biosynthesis.
 SQ SEQUENCE 258 AA; 26815 MW; B2A31BF4B5C6428A CRC64;
 Query Match 57.9%; Score 33; DB 1; Length 258;
 Best Local Similarity 87.5%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 EYOILTFARR 11
 ||| ||| ||
 Db 123 EYOILTFARR 130
 RESULT 7
 GIL2_ENTHI STANDARD; PRT; 288 AA.
 AC 003077;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE GALACTOSE-INHIBITABLE LECTIN 35 KDA SUBUNIT PRECURSOR.
 OS Entamoeba histolytica.
 OC Eukaryota; Entamoebidae; Entamoeba.
 OX NCBI_Taxid=5759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93063055; PubMed=1435873;
 RA Tannich E., Ebert F., Horstmann R.D.;
 RT "Molecular cloning of cDNA and genomic sequences coding for the 35-
 kilodalton subunit of the galactose-inhibitable lectin of pathogenic
 Entamoeba histolytica."
 RL Mol. Biochem. Parasitol. 55:225-227(1992).
 RN [2]
 RP SEQUENCE OF 11-288 FROM N.A.
 RX MEDLINE=94043255; PubMed=8226970;
 RA McCoy J.J., Mann B.J., Vedvick T.S., Pak Y., Helmark D.B.,
 RA Petri W.A. Jr.;
 RT "Structural analysis of the light subunit of the Entamoeba
 histolytica galactose-specific adherence lectin."
 RL J. Biol. Chem. 268:24223-24231(1993).
 CC -1- FUNCTION: LECTIN THAT BINDS GALACTOSE.
 CC -1- SUBUNIT: HETERODIMER OF A HEAVY (170 KDA) AND A LIGHT SUBUNIT
 (35 KDA) LINKED BY DISULFIDE BONDS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL: M96024; AAA29099.1;
 DR EMBL: I06065; AAA16054.1;
 DR PIR: A48452; A48452.
 DR Lectin; Signal.
 KM SIGNAL 1 13
 FT CHAIN 14 288
 FT FT
 FT CONFLICT 11 11 GALACTOSE-INHIBITABLE LECTIN 35 KDA
 FT CONFLICT 64 64 SUBUNIT
 FT CONFLICT 167 167 S -> A (IN REF. 2).
 FT CONFLICT 167 167 E -> K (IN REF. 2).
 FT CONFLICT 167 167 E -> G (IN REF. 2).
 SQ SEQUENCE 288 AA; 33667 MW; DECF366D3DARA24B CRC64;
 Query Match 57.9%; Score 33; DB 1; Length 288;
 Best Local Similarity 54.5%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 TSSEYOILTFARR 11
 ||| ||| ||
 Db 263 TSSEYOILTFARR 273
 RESULT 8
 PARB_CAUCR STANDARD; PRT; 293 AA.
 ID PARB_CAUCR
 AC 005190;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CHROMOSOME PARTITIONING PROTEIN PARB.
 GN PARB.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 CC Caulobacter.
 OX NCBI_Taxid=76;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NA1000;
 RX MEDLINE=97207014; PubMed=9054507;
 RA Mohl D.A., Guber J.W.;
 RT "Cell cycle-dependent polar localization of chromosome partitioning

```

DR InterPro: IPR001714; -
DR InterPro: IPR002467; -
DR Pfam: PF00557; Peptidase_M24; 1.
DR PRINTS: PR00599; MAPEPTIDASE.
DR PROSITE: PS00680; MAP_1; 1.
KM Hypothetical protein: Hydrolase; Aminopeptidase; Cobalt.
FT METAL 147 147 COBALT (BY SIMILARITY).
FT METAL 158 158 COBALT (BY SIMILARITY).
FT METAL 221 221 COBALT (BY SIMILARITY).
FT METAL 255 255 COBALT (BY SIMILARITY).
FT METAL 287 287 COBALT (BY SIMILARITY).
SQ SEQUENCE 305 AA; 33487 MW; 69CCB51DD62EC0BE CRC64;

Query Match 57.9%; Score 33; DB 1; Length 305;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSSEQIITARR 12
   1 1 1 1 1 1 1 1 1 1 1 1
Db 292 TATGYELTDDR 303

RESULT 10
GBBS_HUMAN STANDARD: PRT: 353 AA.
ID GBBS_HUMAN
AC O14775;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT 5 (TRANSDUCIN BETA
DE CHAIN 5).
GN GNB5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98269832; PubMed=9606987;
RX Jones P.G., Lombardi S.J., Cockett M.I.;
RT "Cloning and tissue distribution of the human G protein beta 5 cDNA."
RL Biochim. Biophys. Acta 1402:288-291(1998).
CC CC -1 FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE
CC GTPASE ACTIVITY. FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-
CC EFFECTOR INTERACTION.
CC -1 SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC -1 TISSUE SPECIFICITY: EXPRESSED IN MULTIPLE TISSUES.
CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC
CC EMBL; AF017656; AAC63826.1; -
CC MIM; 604447; -
DR HSSP; P04901; 1GP2.
DR InterPro: IPR001632; -
DR InterPro: IPR001680; -
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00319; GPROTEINB.
DR PRINTS; PR00320; GPROTEINBPT.
DR PROSITE; PS00678; WD_REPEATS_1; 3.
DR PROSITE; PS00682; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
DR Transducer; Repeat; WD repeat, Multigene family.
FT REPEAT 61 91 WD 1.

```

FT REPEAT 103 133 WD 2
 FT REPEAT 151 181 WD 3
 FT REPEAT 193 225 WD 4
 FT REPEAT 237 267 WD 5
 FT REPEAT 281 311 WD 6
 FT REPEAT 323 353 WD 7
 SQ SEQUENCE 353 AA: 38759-MW; 982B08C9F2B9976B CRC64;

Query Match 57.9%; Score 33; DB 1; Length 353;
 Best Local Similarity 70.0%; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TSSEYOILTA 10
 1:1:11111
 DB 162 TNSDMOILTA 171

RESULT 11
 GBS5_MOUSE STANDARD; PRT: 353 AA.

AC P54314: 035354;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT 5 (TRANSDUCIN BETA CHAIN 5).
 GN GNB5.
 OS Mus musculus (Mouse), and Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090, 10116;

[1]
 SEQUENCE FROM N.A.
 RC SPECIES-Rat: TISSUE-Brain;
 RX MEDLINE=94350964; PubMed=80711339;
 RA Watson A.J., Katz A., Simon M.I.;
 RT "A fifth member of the mammalian G-protein beta-subunit family. Expression in brain and activation of the beta 2 isotype of phospholipase C."
 RT J. Biol. Chem. 269:22150-22156(1994).
 RN [2]

RP SEQUENCE FROM N.A.
 RC SPECIES-Rat: STRAIN-SPRAGUE-DAWLEY; TISSUE-Brain;
 RL Levey K., Slepak V.Z.;
 RN Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE OF 11-198 FROM N.A.
 RC SPECIES-Rat: STRAIN-SPRAGUE-DAWLEY;
 RX MEDLINE=9628358; PubMed=9622245;
 RA Betty M., Harnish S.W., Rhodes K.J., Cockett M.I.;
 RT "Distribution of heterotrimeric G-protein beta and gamma subunits in the rat brain."
 RT Neuroscience 85:475-486(1998).

CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS), ARE INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-EFFECTOR INTERACTION.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC -1- TISSUE SPECIFICITY: BRAIN AND AT MUCH REDUCED LEVELS IN KIDNEY.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: L34290: AAA93084.1;
 DR EMBL: AF001953; AAB5974.1;

DR EMBL: AF022086; AAB82553.1;
 DR HSSP: P04901; 1GP2.
 DR MGD: MGI:101848; GNB5.
 DR InterPro: IPR001632;
 DR InterPro: IPR001680;
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00319; GPROTEINB.
 DR PROSITE: PS00678; WD_REPEATS_1; 3.
 DR PROSITE: PS50082; WD_REPEATS_2; 6.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Transducer; Repeat; WD repeat; Multigene family.
 FT REPEAT 61 91 WD 1
 FT REPEAT 103 133 WD 2
 FT REPEAT 151 181 WD 3
 FT REPEAT 193 225 WD 4
 FT REPEAT 237 267 WD 5
 FT REPEAT 281 311 WD 6
 FT REPEAT 323 353 WD 7
 SQ SEQUENCE 353 AA: 38731 MW: 30FCF51C125A024D CRC64;

Query Match 57.9%; Score 33; DB 1; Length 353;
 Best Local Similarity 70.0%; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TSSEYOILTA 10
 1:1:11111
 DB 162 TNSDMOILTA 171

RESULT 12
 YV40_CAEEL STANDARD; PRT: 413 AA.

AC 045335;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HYPOTHETICAL 49.0 KDA PROTEIN F32B6.9 IN CHROMOSOME IV.
 GN F32B6.9
 GN Caenorhabditis elegans.
 OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL NZ;
 RA Basham V.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE BESTROPHIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: Z81074; CAB03043.1;
 DR WormPep: F32B6.9; CE09864.
 DR InterPro: IPR000615;
 DR Pfam: PF01062; Worm.family_8; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 413 AA: 48965 MW: A66E9A83C78790B CRC64;

Query Match 57.9%; Score 33; DB 1; Length 413;
 Best Local Similarity 70.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTA 10
 1:1111111

DB 165 TEKEVILLEA 174

RESULT 13

TRPG_YEAST STANDARD: PRT: 484 AA.

AC P00937;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE ANTHRANILATE SYNTHASE COMPONENT II (EC 4.1.3.27) [INCLUDES: GLUTAMINE AMIDOTRANSFERASE; INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE (EC 4.1.1.48) (PRA1)].

GN TRP3 OR YK1211C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84162082; PubMed=6333449;

RA Zalkin H., Paluh J.L., van Cleemput M., Moye W.S., Yanofsky C.;

RT "Nucleotide sequence of Saccharomyces cerevisiae genes TRP2 and TRP3 encoding bifunctional anthranilate synthase: indole-3-glycerol phosphate synthase.";

RT J. Biol. Chem. 259:3985-3992(1984).

RL [2]

RP SEQUENCE FROM N.A.

RA Pohl T.M., Pohl F.M.;

RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX STRAIN=5288C;

RA MEDLINE=95028164; PubMed=7941750;

RT "The complete sequencing of a 24.6 kb segment of yeast chromosome XI identified the known loci URA1, SAC1 and TRP3, and revealed 6 new open reading frames including homologues to the threonine dehydratase, membrane transporters, hydantoinases and the yeast 10:663-679(1994).

RL [4]

RP SEQUENCE OF 1-280 FROM N.A.

RA Aebi M., Furter R., Prantl F., Niederberger P., Huettler R.;

RT "Structure and function of the TRP3 gene of Saccharomyces cerevisiae: analysis of transcription, promoter sequence, and sequence coding for a glutamine amidotransferase.";

RT Curr. Genet. 8:165-172(1984).

RL [5]

RP CATALYTIC ACTIVITY: CHORISMATE + L-GLUTAMINE - ANTHRANILATE + PYRUVATE + L-GLUTAMATE.

CC -1- CATALYTIC ACTIVITY: 1-(2-CARBOXYPHENYLAMINO)-1-DEOXY-D-RIBULOSE 5-PHOSPHATE - 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE + CO(2) + H(2)O.

CC -1- PATHWAY: FIRST AND FOURTH STEPS IN BIOSYNTHESIS OF TRYPTOPHAN.

CC -1- SUBUNIT: Tetramer of two components I and two components II.

CC -1- INDUCTION: BY TRYPTOPHAN STARVATION.

CC -1- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE USING AMMONIA RATHER THAN GLUTAMINE, WHEREAS COMPONENT II PROVIDES GLUTAMINE AMIDOTRANSFERASE ACTIVITY.

CC -1- MISCELLANEOUS: YEAST COMPONENT II CARBOXYL END ALSO HAS INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE ACTIVITY.

CC -1- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-1 GLUTAMINE AMIDOTRANSFERASES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: K01386; AAA35176.1; -

CC DR EMBL: X75951; CAA53562.1; -

DR EMBL: Z28211; CAA82056.1; -

DR EMBL: M36300; AAA34450.1; -

DR PIR: S38049; NMBY2.

DR PIR: S44327; S44327.

DR HSSP: P00909; IPTI.

DR YEPD: 4596; -

DR SGD: S0001694; TRP3.

DR InterPro: IPR000991; -

DR InterPro: IPR001468; -

DR InterPro: IPR002385; -

DR Pfam: PF00117; GATase; 1.

DR Pfam: PF00218; IGPS; 1.

DR PRINTS: PR00096; GATASE.

DR PRINTS: PR00097; ANTSNTHASEIT.

DR PROSITE: PS00442; GATASE_TYPE_1; 1.

DR PROSITE: PS00614; IGPS; 1.

KW Trypophan biosynthesis; Transferase; Lyase; Multifunctional enzyme; Glutamine amidotransferase.

FT DOMAIN 1 214

FT ACT_SITE 92 484

FT ACT_SITE 92 92

FT ACT_SITE 181 181

FT ACT_SITE 183 183

FT CONFLICT 32 32

FT CONFLICT 63 65

FT CONFLICT 129 129

FT CONFLICT 170 170

FT CONFLICT 236 236

SO SEQUENCE 484 AA; 53489 MW; 34EF65E829279C1F CRC64;

Query Match 57.98; Score 33; DB 1; Length 484;

Best Local Similarity 77.88; Pred. NO. 29;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 348 SKYOLLEA 356

QY 3 SEXOILAR 11

DB 1:||||| 11

DB 348 SKYOLLEA 356

RESULT 14

DOPO_RAT STANDARD: PRT: 620 AA.

AC 005754;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE DOPAMINE BETA-MONOOXYGENASE PRECURSOR (EC 1.14.17.1) (DOPAMINE BETA-HYDROXYLASE) (DBH).

GN DBH.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90219125; PubMed=2325165;

RA McMahon A., Geertman R., Sabban E.L.;

RT "Rat dopamine beta-hydroxylase: molecular cloning and characterization of the cDNA and regulation of the mRNA by reserpine.";

RL J. Neurosci. Res. 25:395-404(1990).

CC -1- FUNCTION: CONVERSION OF DOPAMINE TO NORADRENALINE.

CC -1- CATALYTIC ACTIVITY: 3,4-DIHYDROXYPHENETHYLAMINE + ASCORBATE + O(2) - NORADRENALINE + DEHYDROASCORBATE + H(2)O.

CC -1- COFACTOR: POO, COPPER, AND ASCORBATE.

CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.

CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- SUBCELLULAR LOCATION: EXISTS AS BOTH A SOLUBLE AND MEMBRANE-BOUND FORM IN CHROMAFFIN GRANULES (THE MEMBRANE-BOUND FORM MAY BE ANCHORED BY AN UNCLEAVED SIGNAL SEQUENCE).

CC -1- TISSUE SPECIFICITY: CHROMAFFIN GRANULES OF THE ADRENAL MEDULLA AND SYNAPTIC VESICLES OF THE SYMPATHETIC NERVOUS SYSTEM.

CC -1- SIMILARITY: BELONGS TO THE COPPER TYPE II, ASCORBATE-DEPENDENT
CC MONOOXYGENASE FAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L12407; AAA41091.1;
CC InterPro: IPR000323;
CC DR Pfam: PF01082; Cu2_monooxygen; 1
CC PRINTS: PR00767; DBMONOXGNASE
CC PROSITE: PS00084; CU2_MONOOXYGENASE_1; 1. FALSE_NEG.
CC PROSITE: PS00085; CU2_MONOOXYGENASE_2; FALSE_NEG.
CC Catecholamine biosynthesis; oxidoreductase; Monooxygenase; Copper;
CC Vitamin C; PQQ; Glycoprotein; Membrane; Signal; Phosphorylation.
CC SIGNAL 1 42
CC CHAIN 43 620 POTENTIAL.
CC FT ACT_SITE 233 233 POTENTIAL.
CC FT BINDING 415 415 POTENTIAL.
CC FT MOD_RES 349 349 COPPER (POTENTIAL).
CC FT CARBOHYD 67 67 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 274 274 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 475 475 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 569 569 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 587 587 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 587 587 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 620 AA; 69875 MW; FFE26C84ED5776 CRC64;

Query Match 57.98; Score 33; DB 1; Length 620;
Best Local Similarity 50.0%; Pred. NO. 38;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSSEYOILTARR 12
DB 130 THDIYQLLAQR 141

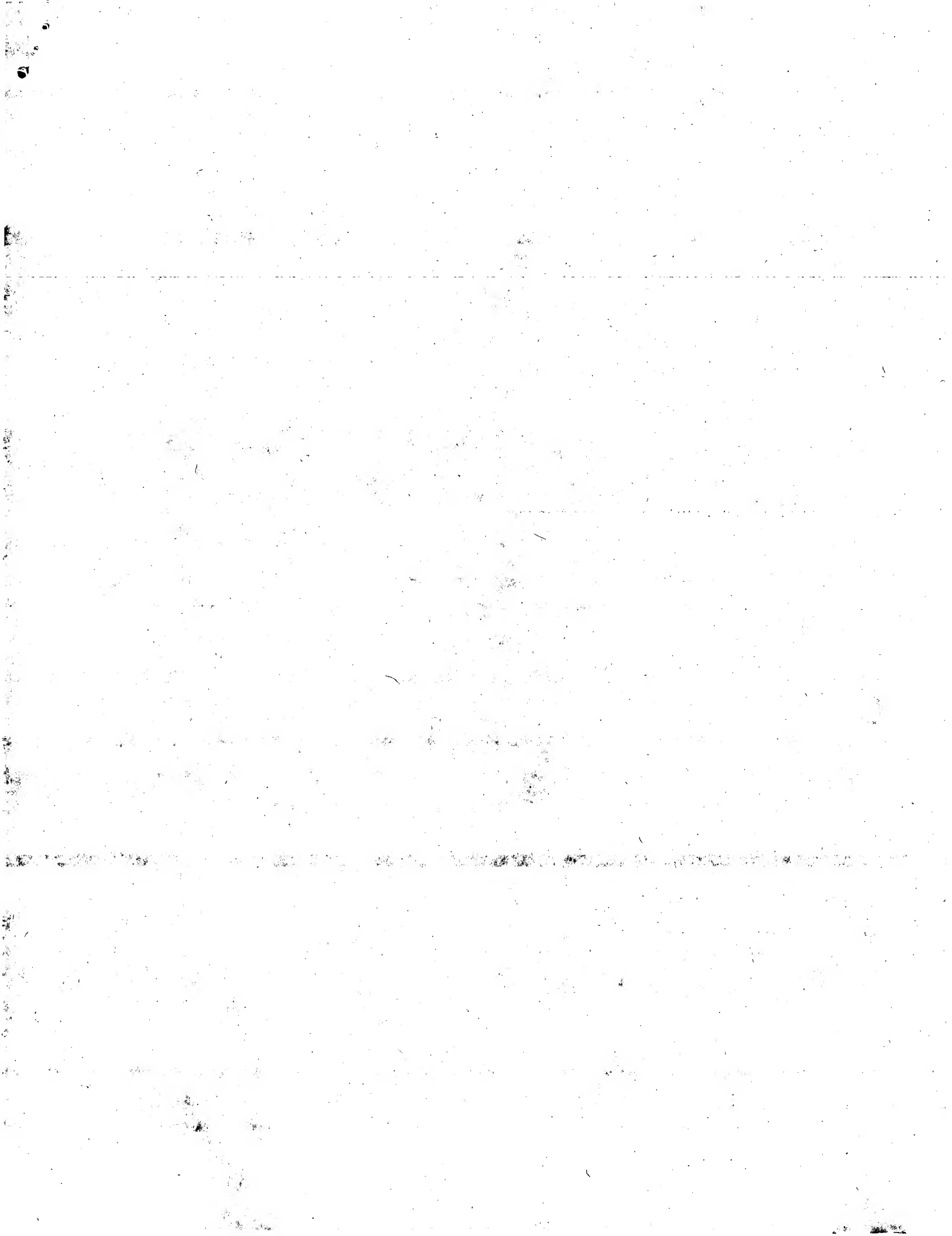
RESULT 15
CJBA_BACUH STANDARD; PRT; 682 AA.
AC 086170; 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PESTICIDIAL CRYSTAL PROTEIN CRY19A (INSECTICIDAL DELTA-ENDOTOXIN
DE CRY19X(A)) (CRYSTALLINE ENTOMOCIDAL PROTOXIN) (78 KDA CRYSTAL
DE PROTEIN)
GN CRY19BA OR CRY19X(A).
OS Bacillus thuringiensis (subsp. higo).
OG Plasmid.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=132266;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98369730; PubMed=9704107;
RA Hwang S.H., Saitoh H., Mizuki E., Higuchi K., Ohba M.;
RT "A novel class of mosquitoicidal delta-endotoxin, Cry19B, encoded by a
RT Bacillus thuringiensis serovar higo gene";
RL Syst. Appl. Microbiol. 21:179-184(1998).
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF MOSQUITOS.
CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT.

CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
CC N-TERMINUS.
CC
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D88381; BAA32397.1;
CC DR HSSP: P07130; IDIC.
CC DR InterPro: IPR001178;
CC DR Pfam: PF00555; endotoxin; 1.
CC DR Toxin; Sporulation; Plasmid.
CC SQ SEQUENCE 682 AA; 78490 MW; 5351EA63E2B042F7 CRC64;

Query Match 57.98; Score 33; DB 1; Length 682;
Best Local Similarity 60.0%; Pred. NO. 42;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SEYOILTARR 12
DB 8 NEYELDAKR 17

Search completed: June 4, 2001, 12:24:06
Job time: 568 sec



AC 092495;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR TYPE I (FC GAMMA RIB-FC GAMMA RECEPTOR).
 GN CD64 OR FC<GAMMA>RIB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Benech P.D., Sasstry K.N., Iyer R.R., Eichbaum Q.G., Raveh D.,
 Ezekowitz A.B.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-224 FROM N.A.
 RX MEDLINE=93018827; PubMed=1402657;
 RA Benech P.D., Sasstry K., Iyer R.R., Eichbaum Q.G., Raveh D.P.,
 Ezekowitz R.A.;
 RT Definition of interferon gamma-response elements in a novel human FC
 gamma receptor gene (Fc gamma RIB) and characterization of the gene
 structure.";
 RT J. Exp. Med. 176:1115-1123(1992).
 RL EMBL: M91555; AAA58414.1;
 DR EMBL: M91550; AAA58414.1; JOINED.
 DR EMBL: M91551; AAA58414.1; JOINED.
 DR EMBL: M91552; AAA58414.1; JOINED.
 DR EMBL: M91553; AAA58414.1; JOINED.
 DR EMBL: M91554; AAA58414.1; JOINED.
 DR EMBL: S45709; AAD13842.1; JOINED.
 DR EMBL: S45707; AAD13842.1; JOINED.
 DR EMBL: S45708; AAD13842.1; JOINED.
 DR EMBL: S45704; AAD13842.1; JOINED.
 DR EMBL: S45705; AAD13842.1; JOINED.
 DR HSSP: P12319; IALT.
 DR INTERPRO: IPR003006;
 DR PRAM: PR000477; 19; 3.
 DR PRODOM: PD002534; 1.
 SQ SEQUENCE 375 AA; 42881 MW; A84D464C7DD0F91 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 57; DB 4; Length 375;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TSSEYQILTRR 12
 DB 242 TSSEYQILTRR 253

RESULT 3
 ID Q9LPH0 PRELIMINARY; PRT: 390 AA.
 AC Q9LPH0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE T3F20.14 PROTEIN.
 GN T3F20.14.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lee J.M., Vaysberg M., Sakano H., Lenz C., Liu S.X., Pham P.,
 Toriumi M., Yu G., Chin C., Choi J., Choi E., Chung M., Gonzalez A.,
 Hong B., Liu A., Altai H., Brooks S., Buehler E., Chao Q., Conn L.,
 Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,

RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
 Ecker J.R., Federspiel N.A., Theologis A.;
 RT "The sequence of BAC T3F20 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC018748; AAF78435.1;
 SQ SEQUENCE 390 AA; 45482 MW; 83E3BCF17110496F CRC64;

Query Match
 Best Local Similarity 64.9%; Score 37; DB 10; Length 390;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSEYQILT 9
 DB 206 SSEYQILT 213

RESULT 4
 ID Q9MZT0 PRELIMINARY; PRT: 349 AA.
 AC Q9MZT0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE FC GAMMA RECEPTOR I.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yan Y., Li X., Wang A., Zhang G.;
 RT "Molecular cloning and identification of full-length cDNA encoding
 RT high affinity FC receptor for bovine IgG (Fc gamma RI).";
 RL Vet. Immunol. Immunopathol. 75:151-159(2000).
 DR EMBL: AF162866; AAF80477.1;
 KW Receptor.
 SQ SEQUENCE 349 AA; 39608 MW; DDB77B2EF9408C02 CRC64;

Query Match
 Best Local Similarity 63.2%; Score 36; DB 6; Length 349;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 TSSEYQILTRR 12
 DB 241 TSSEYQILTRR 252

RESULT 5
 ID Q9L5R1 PRELIMINARY; PRT: 791 AA.
 AC Q9L5R1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN.
 GN R0070.
 OS Salmonella typhi.
 OC Plasmid R27.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OC NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
 Grobeck E., Rose D.J., Taylor D.E.;
 RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
 RT from Salmonella typhi that is temperature sensitive for transfer.";
 RL Nucleic Acids Res. 28:2177-2186(2000).
 DR EMBL: AF250878; AAF69908.1;
 KW Plasmid.

SQ SEQUENCE 791 AA; 89568 MW; 46D6F5751A1C7D67 CRC64;
 Query Match 63.2%; Score 36; DB 2; Length 791;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 TSSEYOILTRAR 12
 DB 632 TDSOYRLVARK 643
 RESULT 6
 ID 033906 PRELIMINARY; PRT; 2004 AA.
 AC 033906;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HYPOHETICAL 216.9 KDA PROTEIN.
 OS Shewanella sp. SCRC-2738.
 OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
 OC Shewanella.
 OX NCBI_TaxId=53560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SCRC-2738;
 RX MEDLINE=97419510; PubMed=9274025;
 RA Takeyama H., Takeda D., Yazawa K., Yamada A., Matsunaga T.;
 RA "Expression of the elcosapentaenoic acid synthesis gene cluster from
 RT Shewanella sp. in a transgenic marine cyanobacterium, Synechococcus
 RT sp."
 RL Microbiology 143:0-0(0).
 DR EMBL: U73935; AAB1125.1; -.
 DR HSSP: P18391; IMKA.
 DR INTERPRO: IPR000794; -.
 DR INTERPRO: IPR001143; -.
 DR INTERPRO: IPR002203; -.
 DR PFAM: PF00109; ketoacyl-synt. 1.
 DR PFAM: PF01377; thioester-denyd; 2.
 DR PROSITE: PS00606; B_KETOACYL-SYNTHASE; UNKNOWN_1.
 DR PROSITE: PS00881; PROTEIN_SPLICING; UNKNOWN_1.
 DR PRODOM: PD013327; -; 2.
 DR Hypothetical protein.
 KM SEQUENCE 2004 AA; 216919 MW; 293CB90F9B53EE2 CRC64;
 Query Match 72.2%; Score 36; DB 2; Length 2004;
 Best Local Similarity 72.7%; Pred. No. 1,3e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 1 TSSEYOILTRAR 11
 DB 905 TASEOQLTRAR 915
 RESULT 7
 ID 09VWA3 PRELIMINARY; PRT; 104 AA.
 AC 09VWA3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG1314 PROTEIN.
 GN CG1314.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernick B.P., Bhandari D., Bolintsov S.,
 RA Borkova D., Botchan M.R., Bouck B., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Canter A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Raine K., Remington K., Saunders R.D.C., Scheelke F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Styksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003514; AAF49045.1; -.
 DR FLYBASE: FBgn0036960; CG13814.
 SQ SEQUENCE 104 AA; 11491 MW; 2620F15C7CBFB431 CRC64;
 Query Match 61.4%; Score 35; DB 5; Length 104;
 Best Local Similarity 60.0%; Pred. No. 8.9;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 3 SEYOILTRAR 12
 DB 90 SEYHLISARR 99
 RESULT 8
 ID 018337 PRELIMINARY; PRT; 197 AA.
 AC 018337;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE COSMTD C31H1.
 GN C31H1.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94150718; PubMed=7906396;
 RX Willson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Tillery-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.".
 RT Nature 368:32-38(1994).
 RN
 RP SEQUENCE FROM N.A.
 RL [2]
 RN Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RN SEQUENCE FROM N.A.
 RL Waterston R.,
 RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U42848; AAA83606.1; -
 DR INTERPRO: IPR000215; -
 DR PROSITE: PS00284; SERPIN; UNKNOWN-1.
 SQ SEQUENCE 197 AA; 22503 MW; 00A5A7EB028E20F4 CRC64;

Query Match 61.4%; Score 35; DB 5; Length 197;
 Best Local Similarity 58.3%; Pred. No. 18;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 TSSEYOILTAR 12
 |||||:|:|:|
 Db 23 TSSEIOLLSMKR 34

RESULT 9
 ID 09W458 PRELIMINARY; PRT; 553 AA.
 AC 09W458; (PRELIMINARY, 13, Created)
 DT 01-MAY-2000 (TREMELREL, 13, Last sequence update)
 DT 01-MAY-2000 (TREMELREL, 13, Last sequence update)
 DE CG12236 PROTEIN.
 GN CG12236
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroides; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blake J.R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferrara C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris W.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jafali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Liang Y., Lin X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster.".
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003435; AAF46100.1; -
 DR FLYBASE: FBgn0029822; CG12236.
 DR INTERPRO: IPR000210; -
 DR INTERPRO: IPR000822; -
 DR PFAM: PF00096; ZF-C2H2; 2.
 DR PFAM: PF00651; BTB; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 2.
 SQ SEQUENCE 553 AA; 60801 MW; EB3019D282B19092 CRC64;

Query Match 61.4%; Score 35; DB 5; Length 553;
 Best Local Similarity 87.5%; Pred. No. 53;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSEYOILT 9
 |||||:|:|:|
 Db 363 SSEYEILT 370

RESULT 10
 ID 09WU16 PRELIMINARY; PRT; 129 AA.
 AC 09WU16; (PRELIMINARY, 12, Created)
 DT 01-NOV-1999 (TREMELREL, 12, Last sequence update)
 DT 01-NOV-1999 (TREMELREL, 12, Last sequence update)
 DE NEUROTRYPHIN-3 PRECURSOR (FRAGMENT).
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 NCBI_TaxID=10036;
 [1]
 RP SEQUENCE FROM N.A.
 RA Jansen H.T., Lehman M.N., Stevens P.J.;
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF104240; AAD21011.1; -
 DR HSSP: P20783; 188K.
 DR INTERPRO: IPR002072; -
 DR PFAM: PF00243; NGE; 1.
 FT NON TER 1 1
 FT NON TER 1 1
 SQ SEQUENCE 129 AA; 14690 MW; 43C825E6651EFDF3 CRC64;

Query Match 59.6%; Score 34; DB 11; Length 129;
 Best Local Similarity 63.6%; Pred. No. 18;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 TSSEYOILTAR 11
 |||||:|:|:|
 Db 41 TRSEFOPTAR 51

RESULT 11
 ID 09SZR4 PRELIMINARY; PRT; 277 AA.

AC 0952R4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HYPOTHEICAL 32.0 KDA PROTEIN.
 GN F27B13.170 OR AT4G29930.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Bancroft I.,
 RA Mewes H.W., Mayer K.F.X., Schueller C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Zimmermann W., Gruenisen A., Wambutt R., Kalicki J., Woldmann P.,
 RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL050352; CAB43668.1;
 DR EMBL: AL161575; CAB79751.1;
 DR INTERPRO: IPR001092;
 DR INTERPRO: IPR003015;
 DR PIRAM: PF00010; HLH; 1
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 277 AA; 31987 MW; 2E85E0FE320F073F CRC64;

Query Match 59.6%; Score 34; DB 10; Length 277;
 Best Local Similarity 58.3%; Pred. No. 41;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
 Db 257 TSTNYKILNASR 268
 RESULT 12
 ID 09P5K7 PRELIMINARY; PRT; 863 AA.
 AC 09P5K7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE RELATED TO GLUCAN 1, 3-BETA-GLUCOSIDASE PRECURSOR PROTEIN.
 GN B23L21.220.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
 NCBI_TaxID=5141;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohnsels J., Brandt P., Fartmann B., Holland R.,
 RA Nakamura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL356172; CAB91690.1;
 SQ SEQUENCE 863 AA; 98928 MW; EC64BADA28708040 CRC64;

Query Match 59.6%; Score 34; DB 3; Length 863;
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
 Db 228 SDEYXYLTAKK 239

RESULT 13
 ID 09LTX7 PRELIMINARY; PRT; 1080 AA.
 AC 09LTX7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GB1AF56406.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA SATO S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RC STRAIN=COLUMBIA;
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty pl and TAC
 RT clones."
 RL DNA Res. 7:31-63(2000).
 DR EMBL: AB024032; BAA97014.1;
 SQ SEQUENCE 1080 AA; 121430 MW; CSC1B2D9173215C6 CRC64;

Query Match 59.6%; Score 34; DB 10; Length 1080;
 Best Local Similarity 58.3%; Pred. No. 1.8e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 TSSEYOILTARR 12
 Db 133 TDSYMITTLR 144

RESULT 14
 ID 09QXL1 PRELIMINARY; PRT; 1668 AA.
 AC 09QXL1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE KIF21B.
 GN KIF21B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99242603; PubMed=10225949;
 RA Marszalek J.R., Weiner J.A., Farlow S.J., Chun J., Goldstein L.S.;
 RT "Novel dendritic kinesin sorting identified by different process
 RT targeting of two related kinesins: KIF21A and KIF21B."
 RL J. Cell Biol. 145:469-479(1999).
 DR EMBL: AF202893; AAF17084.1;
 DR HSSP: P17119; 3KAR.
 DR MGD: MGI:109234; Kif21b.
 DR INTERPRO: IPR001680;
 DR INTERPRO: IPR001752;
 DR PIRAM: PF00225; kinesin_1.
 DR PIRAM: PF00400; WD40; 7.
 DR PRINTS: PR00380; KINESINHEAVY.
 DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; UNKNOWN_1.

DR PROSITE: PS00678; WD_REPEATS; UNKNOWN.1.
DR PROSITE: PS00667; KINESIN_MOTOR_DOMAIN2; 1.
SQ SEQUENCE 1668 AA: 186162 MW: 371452867C9A57 CRC64.

Query Match	59.6%;	Score 34;	DB 11;	Length 1668;
Best Local Similarity	54.5%;	Pred. No. 2.8e+02;		
Matches	6;	Conservative	3;	Mismatches 2;
				Indels 0;
				Gaps 0.

QY	1	TSSEYÖILTAR	11
		1:11:111:	
Db	259	TGTEYEFTLTAK	269

RESULT	15			
09NPM7		PRELIMINARY:	PRT:	271 AA.
ID	09NPM7			
AC	09NPM7			
DT	01-OCT-2000 (Tremblrel, 15, Created)			
DR	01-OCT-2000 (Tremblrel, 15, Last sequence update)			
DT	01-OCT-2000 (Tremblrel, 15, Last annotation update)			
DE	HYPOTHETICAL 30.8 KDA PROTEIN (FRAGMENT).			
GN	DKFZP62M186.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=MELANOMA (MEMO CELL LINE);			
RA	Bloesch-H.L.N., Boesch M., Brandt P., Newes H.W., Weil B., Wiemann S.;			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, AL359590: CAB94876.1;			
KW	Hypothetical protein;			
FT	NON-ITER 1			
SEQ	SEQUENCE 271 AA: 30764 MM: 011097C174878478 CRC64;			

Query Match	57.9%	Score 33;	DB 4;	Length 271;
Best Local Similarity	70.0%;	Pred. No. 66;		
Matches	7;	Mismatches	2;	Indels 0; Gaps 0;

QY	1	TSSEYQILTA	10
		:	
Db	117	TRAIFYQILLA	126

Search completed: June 4, 2001, 12:23:27
Job time: 590 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:30 ; Search time 106.14 Seconds
(without alignments)
6.463 Million cell updates/sec

Title: US-09-284-107-26
Perfect score: 64
Sequence: 1 EDSGLYCEAF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues
Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq.0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	12	19 W60558	Oligopeptide from
2	57	89.1	261	20 Y33183	Human sfc-gammari
3	57	89.1	344	13 R22549	Human macrophage-s
4	57	89.1	344	17 R91439	Human FCRI (CDNA c
5	57	89.1	344	21 Y96183	Human macrophage-s
6	57	89.1	374	13 R20811	Human macrophage-s
7	57	89.1	374	13 R22550	Human macrophage-s
8	57	89.1	374	17 R91438	Human FCRI (CDNA c
9	57	89.1	374	17 W00859	Human FCRI (CDNA c
10	57	89.1	374	21 Y96134	Human macrophage-s
11	57	89.1	374	21 Y96226	Human high affinity

12	57	89.1	399	21 B43683	Human cancer assoc
13	48	75.0	374	19 W80448	Human Fc receptor
14	48	75.0	374	19 W97833	Human Fc receptor
15	48	75.0	374	19 W97834	Human Fc receptor
16	48	75.0	377	20 W86195	Human Fc receptor
17	48	75.0	1018	18 W06485	Rat contactin 11ga
18	48	75.0	1125	19 W52288	Rattus norvegicus
19	48	75.0	1139	19 W37779	Rattus norvegicus
20	48	75.0	1242	19 W52287	Rattus norvegicus
21	48	75.0	1251	19 W37778	Rattus norvegicus
22	47	73.4	1447	16 R68553	Deleted in colorec
23	47	73.4	1447	20 Y33498	Human DCC protein.
24	47	73.4	1447	22 B50693	Human UNC-40 prote
25	47	73.4	1728	12 R13144	Deleted in Colorec
26	46	71.9	150	20 Y02365	Polyptide identl
27	46	71.9	234	19 W74886	Human secreted pro
28	46	71.9	234	21 Y87334	Human signal pepti
29	46	71.9	410	12 R12428	Hybrid Fc(gamma)RI
30	46	71.9	1018	15 R63759	Human contactin (E
31	46	71.9	1018	17 R87028	Human contactin.
32	44	68.8	561	17 R91065	Rat Interleukin-1
33	44	68.8	582	17 R92256	Neural cell adhesi
34	44	68.8	761	17 R92255	Neural cell adhesi
35	44	68.8	4412	21 Y53666	Sequence q1/101742
36	43	67.2	272	17 R92254	Neural cell adhesi
37	43	67.2	333	21 B12313	Human secreted pro
38	43	67.2	398	21 Y96307	Human IGFAM-19 Imm
39	43	67.2	517	21 Y32390	Herpesvirus entry
40	43	67.2	738	20 Y29639	Human VEGF recepto
41	43	67.2	738	21 Y59302	Human VEGF recepto
42	43	67.2	764	19 W69679	KDR protein sequen
43	43	67.2	789	15 R62488	Truncated KLD SVES
44	43	67.2	789	20 W68007	Human soluble vasc
45	43	67.2	828	12 R13905	Rabbit ATHERO-ELAM

ALIGNMENTS

RESULT 1	
ID W60558	standard; peptide; 12 AA.
AC W60558;	
DT 18-AUG-1998	(first entry)
DE Oligopeptide from extracellular domain of CD64.	
KW Extracellular domain; CD64; identification; antibody;	
KM Immunohistochemical; immunofluorescent analysis; detection;	
KW cell transformation; mutation; anti; oncogene.	
OS Synthetic.	
PN W09815833-A1.	
PD 16-APR-1998.	
PF 07-OCT-1997;	97WO-NL00557.
PK 08-OCT-1996;	96BP-0202791.
PA (UYUT-) RIKUSUNIV UTRCHT.	
PI De Krulif CA, Logtenberg T;	
DR WPI; 1998-240964/21.	
XX Identifying peptide(s) binding specifically to protein target - by	
XX expressing on phage surface and testing for binding to immobilised	
PT oligopeptide derived from the target, useful for, e.g. identifying	
PT specific antibodies	

XX Example 1: Page 29, 40pp; English.
 PS
 CC Synthetic oligopeptides W6037-61 are derived from the extracellular
 CC domain of CD64. They were synthesised on 25 polyethylene rods as
 CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
 CC affinity binding to the CD64-covered rods. The specification describes a
 CC method for the identification of a peptide able to bind specifically to a
 CC target protein. The method comprises displaying the peptide on the
 CC surface of a replicable display package, synthesising oligopeptides
 CC derived from the target protein on a solid phase, and testing for binding
 CC between the peptide and oligopeptides. The method is used to screen large
 CC peptide libraries, especially to detect antibodies, or their fragments,
 CC that bind to cell markers or that can differentiate between different
 CC forms of the same protein, including bispecific antibodies that bind to
 CC two non-overlapping epitopes on the same monomeric antigen or two
 CC epitopes on different molecules. The genes/oligonucleotides that encode
 CC selected peptides can be isolated and used for recombinant production of
 CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
 CC immunofluorescent analysis, and also to detect cell transformation caused
 CC by mutation in (anti)oncogenes.
 CC
 SQ Sequence 12 AA;

Query Match 100.0%; Score 64; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00033;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGSLYQCEAAT 12
 |||||
 DB 1 edgslgyceaat 12

RESULT 2

Y33183
 ID Y33183 standard; Protein; 261 AA;

AC Y33183;

DT 15-NOV-1999 (first entry)

DE Human sFc-gammaRI protein fragment.

XX
 KW Fc receptor; Fc-gammaRI; human; FCR; model; three-dimension; 3-D;
 KW atomic coordinate; bioactive compound design; computer-assisted;
 KW drug design; therapy; inhibitor; Fc-gammaR; Fc-epsilonR; IgG; IgE;
 KW tissue damage; hypersensitivity; inflammatory cell recruitment;
 KW inflammatory modulator; Fc-gammaRI; immune function regulation;
 KW anti-inflammatory; immunoprotective; sFc-gammaRI.
 XX
 OS Homo sapiens.
 XX
 PN WO9940117-A1.
 PD 12-AUG-1999.
 PF 04-FEB-1999; 99WO-IB00367.
 PR 11-SEP-1998; 98US-0098994.
 PR 06-FEB-1998; 98US-0073972.
 XX
 PA (ILEX-) ILEXUS PTY LTD.
 PI Baell JB, Epa V, Garrett TJ, Hogarth PM, Matthews BR,
 PI Maxwell KF, McCarthy TD, McKenzie IF, Plesters GA;
 PI Powell MS;
 PS WPI, 1999-539978/45.
 DR
 XX
 PT Three-dimensional structures and models of Fc receptors, useful in
 XX computer-assisted drug design

PS Claim 11: Page 316-317; 326pp; English.

XX
 CC This invention describes a novel model of an Fc-receptor (FcR) protein
 CC representing a three-dimensional (3-D) structure that substantially
 CC conforms to the specified atomic coordinates. Computer model images of
 CC the FcR can be used to design bioactive chemical compounds, e.g.
 CC oligonucleotides, peptides, peptidomimetics and small organic molecules,
 CC by computer-assisted methods of drug design. Therapeutic compositions
 CC that inhibit the activity of Fc-gammaR or Fc-epsilonR can be used to
 CC reduce IgG- or IgE-mediated tissue damage, e.g. IgG- or IgE-mediated
 CC hypersensitivity; recruitment of inflammatory cells or release of
 CC inflammatory mediators. The therapeutic compositions can also be used to
 CC treat disease in animals. The knowledge of the 3-D structure of Fc-gamma
 CC R1a and models of other FcR provides a means for designing and producing
 CC compounds that regulate immune function and inflammation in an animal,
 CC including humans (i.e. structure based drug design). For example,
 CC chemical compounds can be designed to block binding of immunoglobulin to
 CC an Fc receptor protein using various computer programs and models. The
 CC products of the invention have anti-inflammatory and immunoprotective
 CC activity. This sequence represents the human sFc-gammaRI protein
 CC fragment described in the method of the invention.
 CC
 SQ Sequence 261 AA;

Query Match 89.1%; Score 57; DB 20; Length 261;
 Best Local Similarity 91.7%; Pred. No. 0.098;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSLYQCEAAT 12
 |||||
 DB 234 edgslgyceaat 245

RESULT 3

R22549
 ID R22549 standard; Protein; 344 AA.

AC R22549;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FcRI receptor encoded by clone p98.

XX
 KW Rapid immunoselection cloning technique; cell surface antigen;
 KW immunodiagnosis; high affinity receptor.
 XX
 OS Homo sapiens.
 XX
 PN WO9201049-A.
 PD 23-JAN-1992.
 PF 15-JUL-1990; 90WO-US04986.
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO-) GEN HOSPITAL CORP.
 PI Seed B, Aruffo A, Amlot M;
 PI WPI, 1992-056864/07.
 DR N-PSDB; Q21179.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX
 PS Example 10: Page 94a; 160pp; English.

CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p98/X2. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC while clone p98/X2 predicts a Thr residue. At position 58, p90

CC (see Q21180) predicts Val and p98/X2 predicts Leu. Sequences
 CC predicted from all 3 clones show the typical features of a type I
 CC integral membrane protein and include a short hydrophobic signal
 CC sequence, a single 21-residue hydrophobic membrane-spanning domain,
 CC and a short, highly charged cytoplasmic domain. The extracellular
 CC portion contains six potential N-linked glycosylation sites and six
 CC Cys residues distributed among three C2 set Ig-related domains.
 CC A fusion protein of FCRI and a receptor ligand will be helpful to
 CC increase the potency of antibodies in therapy.

SO Sequence 344 AA:

Query Match 89.1%; Score 57; DB 13; Length 344;

Best Local Similarity 91.7%; Pred. No. 0.13;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDSGLYOCEAAT 12
 ||||| |||||
 Db 253 edsglylweaat 264

RESULT 4

ID R91439 standard; Protein; 344 AA.

AC R91439;

DT 30-OCT-1996 (first entry)

DE Human FCRI (CDNA clone p98 product).

KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FCRI; Fc receptor.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 58 /note= "amino acid 58 is Leu in clone p135
 translated product"

PN US5506126-A.

XX 09-APR-1996.

PF 25-FEB-1988; 88US-0160416.

XX 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 18-OCT-1993; 93US-0139273.

PA (GEHO) GEN HOSPITAL CORP.

PI Arufio A, Seed B;

XX WPI; 1996-200279/20.

DR N-PSDB; T14718.

XX Cloning of cDNA encoding cell surface antigen - useful for isolation
 of diagnostic and therapeutic proteins

FT Example 10; Column 55-56; 79pp; English.

PS The amino acid sequence (R91439) of human FCRI was detd. from a

CC cDNA clone, p98 (T14718), obtd. from a cDNA library using an

CC immunoselection cloning method. FCRI is a high affinity receptor

CC for the Fc portion of IgG, normally located on cell surfaces of

CC macrophages. 2 Other cDNA clones, p135 (T14717) and p90 (T14719),

CC coded for a variants of the FCRI sequence; the C-terminal sequence

CC and p90 products (see also R91438 and W00859). A fusion protein
 CC of FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.

SO Sequence 344 AA:

Query Match 89.1%; Score 57; DB 17; Length 344;

Best Local Similarity 91.7%; Pred. No. 0.13;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDSGLYOCEAAT 12
 ||||| |||||
 Db 253 edsglylweaat 264

RESULT 5

ID Y96183 standard; Protein; 344 AA.

AC Y96183;

DT 19-DEC-2000 (first entry)

DE Human macrophage-specific FCRI.

KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
 KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 15 /note= "encoded by GCG"

FT Misc-difference 38 /note= "encoded by ACC"

FT Misc-difference 50 /note= "encoded by GCG"

FT Misc-difference 51 /note= "encoded by ACC"

FT Misc-difference 55 /note= "encoded by ACC"

FT Misc-difference 56 /note= "encoded by CAC"

FT Misc-difference 60 /note= "encoded by TCC"

FT Misc-difference 64 /note= "encoded by CCC"

FT Misc-difference 82 /note= "encoded by CAG"

FT Misc-difference 82 /note= "encoded by CAA"

FT Misc-difference 116 /note= "encoded by GAA"

FT Misc-difference 117 /note= "encoded by GCA"

FT Misc-difference 122 /note= "encoded by TTC"

FT Misc-difference 123 /note= "encoded by ACC"

FT Misc-difference 126 /note= "encoded by CCG"

FT Misc-difference 129 /note= "encoded by CAT"

FT Misc-difference 134 /note= "encoded by AAT"

FT Misc-difference 136 /note= "encoded by GTT"

FT Misc-difference 139 /note= "encoded by CCA"

FT Misc-difference 140 /note= "encoded by AAT"

FT Misc-difference 213 /note= "encoded by AAT"

```

FT  Misc-difference 216 /note= "encoded by CAA"
FT  Misc-difference 220 /note= "encoded by TTC"
FT  Misc-difference 268 /note= "encoded by CGT"
FT  Misc-difference 305 /note= "encoded by AAT"
FT  Misc-difference 306 /note= "encoded by GTG"
FT  Misc-difference 332 /note= "encoded by AAC"
FT  Misc-difference 333 /note= "encoded by GGT"
FT  Misc-difference 338 /note= "encoded by GGC"
FT  Misc-difference 338 /note= "encoded by CCG"
XX  US6111093-A.
XX  29-AUG-2000.
XX  28-OCT-1998; 98US-0181612.
XX  01-DEC-1992; 92US-0983647.
XX  25-FEB-1988; 88US-0160416.
XX  13-JUL-1989; 89US-0379076.
XX  23-MAR-1990; 90US-0498809.
XX  13-JUL-1990; 90US-0553759.
XX  (GENO ) GEN HOSPITAL CORP.
XX  Stamenkovic I, Seed B;
XX  WPI; 2000-586382/55.
XX  N-PSDB; A50631.
XX  Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
XX  useful for immunodiagnosis and immunotherapy of immune-mediated,
XX  infections or disorders, e.g. asthma, immune-complex disease, parasitic
XX  diseases
XX  Example 10; Column 53-55; 75pp; English.
XX  The present sequence is that of a human macrophage specific FCRI,
XX  as deduced from cDNA clone p135 (see A50631), which was isolated
XX  from a cDNA library expressed in COS cells using a novel method of
XX  the invention designed to isolate CSA nucleic acids. The method is
XX  based upon transient expression of a CSA in eukaryotic cells and
XX  physical selection of cells expressing the antigen by adhesion to
XX  (panning on) an antibody-coated substrate such as a culture dish.
XX  CSA nucleic acids isolated by the method of the invention, and the
XX  proteins they encode, are useful for immunodiagnosis and
XX  immunotherapeutic applications, including the diagnosis and
XX  treatment of immune-mediated infections, diseases, and disorders in
XX  animals, including humans. These disorders include asthma,
XX  immune-complex disease, amyloidosis, parasitic diseases or multiple
XX  sclerosis. FCRI is a high affinity receptor for the Fc portion of
XX  IgG, normally located on the cell surfaces of macrophages. The
XX  ability to interfere with such bonding, or to cause it to occur on
XX  surfaces other than macrophages, is useful in therapy. A fusion
XX  protein of FCRI and a receptor ligand will be helpful to increase
XX  the potencies of antibodies in therapy.
XX  Sequence 344 AA;

```

```

Query Match 89.1%; Score 57; DB 21; Length 344;
Best Local Similarity 91.7%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EDGGLYCEAAT 12
| | | | | | | | | |
XX

```

```

Db 253 edsglywceat 264
RESULT 6
ID R20811
XX R20811 standard; Protein; 374 AA.
XX
XX R20811;
XX
XX 21-MAY-1992 (first entry)
XX
XX Human macrophage-specific FCRI receptor encoded by clone p135.
XX Rapid immunoselection cloning technique; cell surface antigen;
XX Immunodiagnosis; high affinity receptor.
XX Homo sapiens.
XX OS
XX PN MO9201049-A.
XX PD 23-JAN-1992.
XX PF 15-JUL-1990; 90MO-US04986.
XX PR 13-JUL-1990; 90US-0553759.
XX PA (GENO-) GEN HOSPITAL CORP.
XX PI Seed B, Aruffo A, Amiot M;
XX
XX WPI; 1992-056864/07.
XX N-PSDB; Q21178.
XX
XX New CD53 cell surface antigen and DNA encoding it - for
XX immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
XX Example 10; Page 94a; 160pp; English.
XX
XX This amino acid sequence was predicted from the cDNA sequence of
XX cDNA clone p135. It differs from the sequence predicted from
XX clones p98/X2 and p90 (see Q21179 and Q21180, respectively) at
XX position 25; p135 encodes a Ser residue and the other two clones
XX predict a Thr residue. At position 58, p135 predicts Leu and p90
XX predicts Val. Sequences predicted from all 3 clones show the
XX typical features of a type I integral membrane protein and include
XX a short hydrophobic signal sequence, a single 21-residue
XX hydrophobic membrane-spanning domain, and a short, highly charged
XX cytoplasmic domain. The extracellular portion contains six
XX potential N-linked glycosylation sites and six Cys residues
XX distributed among three C2 set Ig-related domains. A fusion protein
XX of FCRI and a receptor ligand will be helpful to increase the
XX potency of antibodies in therapy.
XX
XX Sequence 374 AA;

```

```

Query Match 89.1%; Score 57; DB 13; Length 374;
Best Local Similarity 91.7%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EDGGLYCEAAT 12
| | | | | | | | | |
Db 253 edsglywceat 264
RESULT 7
R22550
ID R22550 standard; Protein; 374 AA.
XX R22550;
XX
XX 21-MAY-1992 (first entry)
XX

```

DE Human macrophage-specific FCRI receptor encoded by clone p90.
 XX
 KW Rapid immunoselection cloning technique; cell surface antigen;
 KW Immunodiagnosis; high affinity receptor.
 XX
 OS Homo sapiens.
 XX
 PN W09201049-A.
 XX
 PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90MO-US04986.
 XX
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO-) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Amiot M;
 XX
 XX WPI: 1997-056864/07.
 DR N-PSDB; Q21180.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX
 PS Example 10; Page 94a; 160pp; English.
 XX
 CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p90. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC and clone p90 predicts a Thr residue. At position 58, p90 predicts
 CC Val while clones p135 and p98/X2 (see Q21179) both predict Leu.
 CC Sequences predicted from all 3 clones show the typical features of
 CC a type I integral membrane protein and include a short hydrophobic
 CC signal sequence, a single 21-residue hydrophobic membrane-spanning
 CC domain, and a short, highly charged cytoplasmic domain. The
 CC extracellular portion contains six potential N-linked glycosylation
 CC sites and six Cys residues distributed among three C2 set Ig-related
 CC domains. A fusion protein of FCRI and a receptor ligand will be
 CC helpful to increase the potency of antibodies in therapy.
 CC
 CC Sequence 374 AA;
 SO

Query Match 89.1%; Score 57; DB 13; Length 374;
 Best Local Similarity 91.7%; Pred. No. 0.14;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSLGYCEAAT 12
 ||||| |||||
 Db 253 edsglywceaat 264

RESULT 8
 R91438
 ID R91438 standard; Protein; 374 AA.
 AC R91438;
 XX
 DT 30-OCT-1996 (first entry)
 XX
 DE Human FCRI (cDNA clone p135 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FCRI; Fc receptor.
 XX
 OS Homo sapiens.
 XX
 PN Key
 XX Location/Qualifiers
 FT Misc-difference 25 /note="amino acid 25 is Thr in clone p90 and p98
 FT translated products"
 FT Misc-difference 58

FT /note="amino acid 58 is Val in p90 clone
 FT translated product"
 FT
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 XX WPI: 1996-200279/20.
 DR N-PSDB; T14717.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX
 PS Example 10; Column 55-56; 79pp; English.
 XX
 CC The amino acid sequence (R91438) of human FCRI was detd. from a
 CC cDNA clone, p135 (T14717), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Another isolated cDNA clone, p90 (T14719), coded
 CC for a variant (W00859) of the FCRI sequence, and a third clone, p98
 CC (T14718), coded for an FCRI (R91439) having a different C-terminal
 CC sequence. A fusion protein of FCRI and a receptor ligand will
 CC be helpful in increasing the potency of antibodies in therapy.
 CC
 CC Sequence 374 AA;
 SO

Query Match 89.1%; Score 57; DB 17; Length 374;
 Best Local Similarity 91.7%; Pred. No. 0.14;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSLGYCEAAT 12
 ||||| |||||
 Db 253 edsglywceaat 264

RESULT 9
 W00859
 ID W00859 standard; Protein; 374 AA.
 AC W00859;
 XX
 DT 30-OCT-1996 (first entry)
 XX
 DE Human FCRI (cDNA clone p90 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; FCRI; Fc receptor.
 XX
 OS Homo sapiens.
 XX
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.

	Query Match	89.1%	Score 57;	DB 17;	Length 374;
	Best Local Similarity	91.7%;	Pred.	No. 0 14;	
	Matches 11; Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
OY	1 EDSGLYOCEAT 12				
DG	253 edsglywceat 264				

FT	Misc-difference	117	/note=	"encoded by GAA"
FT	Misc-difference	117	/note=	"encoded by GAA"
FT	Misc-difference	122	/note=	"encoded by GCA"
FT	Misc-difference	122	/note=	"encoded by TTC"
FT	Misc-difference	123	/note=	"encoded by ACC"
FT	Misc-difference	126	/note=	"encoded by CCG"
FT	Misc-difference	129	/note=	"encoded by CAT"
FT	Misc-difference	134	/note=	"encoded by AAT"
FT	Misc-difference	136	/note=	"encoded by GAT"
FT	Misc-difference	139	/note=	"encoded by CCA"
FT	Misc-difference	140	/note=	"encoded by AAT"
FT	Misc-difference	213	/note=	"encoded by CAA"
FT	Misc-difference	216	/note=	"encoded by TTC"
FT	Misc-difference	220	/note=	"encoded by CCG"
FT	Misc-difference	268	/note=	"encoded by AAT"
FT	Misc-difference	305	/note=	"encoded by GAT"
FT	Misc-difference	306	/note=	"encoded by CCA"
FT	Misc-difference	332	/note=	"encoded by AAC"
FT	Misc-difference	332	/note=	"encoded by GGT"
XX	US6111093-A.			
PD	29-AUG-2000.			
XX	28-OCT-1998.	9805-0181612.		
XX	01-DEC-1992;	9205-0983647.		
XX	25-FEB-1988;	8805-0160416.		
XX	13-JUL-1989;	8905-0379076.		
XX	23-MAR-1990;	9005-0498809.		
XX	13-JUL-1990;	9005-0535759.		
XX	(GEO) GEN HOSPITAL CORP.			
XX	Stamenkovic I, Seed B;			
XX	WPI: 2000-586382/55.			
XX	N-PSDB; A50592.			
XX	Isolated nucleic acid molecule encoding the CD19 cell surface antigen,			
XX	useful for immunodiagnosis and immunotherapy of immune-mediated			
XX	infections or disorders, e.g. asthma, immune-complex disease, parasitic			
XX	diseases			
XX	Example 10; Column 53-55; 75pp; English.			
XX	The present sequence is that of a human macrophage specific FCRI,			
XX	as deduced from cDNA clone p135 (see A50592), which was isolated			
XX	from a cDNA library expressed in COS cells using a novel method of			
XX	the invention designed to isolate CSA nucleic acids. The method is			
XX	based upon transient expression of a CSA in eukaryotic cells and			
XX	physical selection of cells expressing the antigen by adhesion to			
XX	(panning on) an antibody-coated substrate such as a culture dish.			
XX	CSA nucleic acids isolated by the method of the invention, and the			
XX	proteins they encode, are useful for immunodiagnosis and			
XX	immunotherapeutic applications, including the diagnosis and			
XX	treatment of immune-mediated infections, diseases, and disorders in			
XX	animals, including humans. These disorders include asthma,			
XX	immune-complex disease, amyloidosis, parasitic diseases or multiple			

CC sclerost. FcRI is a high affinity receptor for the Fc portion of
CC IgG, normally located on the cell surfaces of macrophages. The
CC ability to interfere with such bonding, or to cause it to occur on
CC surfaces other than macrophages, is useful in therapy. A fusion
CC protein of FcRI and a receptor ligand will be helpful to increase
CC the potencies of antibodies in therapy.

XX Sequence 374 AA:

Query Match 89.1%; Score 57; DB 21; Length 374;
Best Local Similarity 91.7%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGGLYQCEAAT 12
||||| |||||
Db 253 edsglywceat 264

RESULT 11.

ID Y96226 standard; Protein: 374 AA.

AC Y96226;

DT 11-SEP-2000 (first entry)

DE Human high affinity Fc receptor, Fcgammar1.

XX Human; high affinity Fc receptor; Fcgammar1; immunoglobulin;

KW infection; immune response; CD64; monocyte; macrophage; neutrophil;

KW eosinophil; HIV; IgG; immunosuppressive; antirheumatic; cytostatic;

KW antiinflammatory; AIDS; SLE; multiple myeloma; rheumatoid arthritis;

KW systemic lupus erythematosus; tumour.

XX Homo sapiens.

OS Homo sapiens.

PN EP1006183-A1.

XX 07-JUN-2000.

XX 03-DEC-1996; 98EP-0122969.

XX 03-DEC-1998; 98EP-0122969.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX WPI: 2000-367968/32.

XX N-PSDB: A27466.

XX Novel Fc receptor lacking transmembrane domains, a signal peptide, and

XX glycosylation, useful for diagnosing and treating immune disorders and

XX cancer

XX Disclosure: Page 26-28; 60pp; English.

XX The present sequence is the human high affinity Fc receptor, Fcgammar1.

XX Fcgammar1 is also known as CD64. Fc receptors play an important

XX role in defending the body against infections. First, pathogens are

XX opsonised by serum immunoglobulins. The resulting complex then binds to

XX cells expressing Fc receptors. Fcgammar1 molecules are expressed by

XX monocytes and macrophages, but expression can also be induced on

XX neutrophils and eosinophils. Upon Fc receptor activation, immune effector

XX pathways are activated, leading to immune response. The present sequence

XX may be modified to produce recombinant versions. The recombinant Fc

XX receptor consist only of the extracellular portion of the receptor and

XX are not glycosylated i.e. they do not have transmembrane domains or

XX signal peptides. The recombinant proteins may be used in immunoassays to

XX determine the immune status of patients with chronic diseases of the

XX immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple

XX myeloma (MM), or rheumatoid arthritis. In addition, pharmaceutical

XX compositions containing recombinant proteins may be used to treat or

XX prevent autoimmune diseases, allergies or tumours, especially AIDS.

CC rheumatoid arthritis or MM.

XX Sequence 374 AA:

Query Match. 89.1%; Score 57; DB 21; Length 374;
Best Local Similarity 91.7%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGGLYQCEAAT 12
||||| |||||
Db 253 edsglywceat 264

RESULT 12

ID B43683 standard; Protein: 399 AA.

AC B43683;

DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1128.

XX Human; cancer associated gene; cancer antigen; detection; cancer;

KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;

KW antidiabetic; antistimatic; antirheumatic; antithyroid; antiviral;

KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;

KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;

KW vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;

KW immune disorder; hematopoietic cell disorder; autoimmune disorder;

KW allergic reaction; graft versus host disease; organ rejection;

KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurological disease; drug screening.

XX Homo sapiens.

OS Homo sapiens.

PN WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM,

XX WPI: 2000-587533/55.

XX N-PSDB: C77892.

XX Novel isolated nucleic acids comprising sequences encoding peptides

XX useful for treating or diagnosing e.g. cancer

XX Claim 11; Page 1739-1740; 2352pp; English.

XX C77607 to C78448 encode the human cancer associated proteins given in

XX B43398 to B44239. The proteins can have activities based on the tissues

XX and cells the genes are expressed in. Example of activities include:

XX cytostatic; proliferative; vulnery; immunomodulator; antidiabetic;

XX antistimatic; antirheumatic; antithyroid; antiviral; inflammatory;

XX antithyroid; antiallergic; antibacterial; coagulant; neutropic;

XX neuroprotective; cardiant; thrombolytic; coagulant; neutropic;

XX vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;

XX polypeptides can be used for preventing, treating or ameliorating medical

XX conditions and diagnosing pathological conditions. Polynucleotides,

XX polypeptides, antibodies, agonists and antagonists from the present

XX invention may be used to treat immune disorders by activating or

XX inhibiting the proliferation, differentiation or mobilisation of immune

XX cells, to treat disorders of haematopoietic cells, autoimmune disorders,

XX allergic reactions, graft versus host disease and organ rejection,

XX modulate haemostatic or thrombolytic activity, modulate inflammation,

CC cancers, cardiovascular disorders, neurological disease and bacterial or
 CC viral infections. The peptides, nucleotides, antibodies, agonists and
 CC antagonists may be also be used in drug screens. C78449 to C78457 and
 CC B44240 represent sequences used in the exemplification of the present
 CC invention.

CC Sequence 399 AA;

Query Match 89.1%; Score 57; DB 21; Length 399;
 Best Local Similarity, 91.7%; Pred. No. 0.15;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDSGLYQCEAT 12
 DB 278 edsglywceat 289

RESULT 13
 ID W80448 standard; Protein: 374 AA.

AC W80448;

DT 07-JUN-1999 (first entry)

DE Human Fc receptor I.

XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 XX cloning.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 2 /note= "encoded by TGC"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by CTC"
 FT Misc-difference 60 /note= "encoded by CCC"
 FT Misc-difference 77 /note= "encoded by AAT"
 FT Misc-difference 85 /note= "encoded by TCC"
 FT Misc-difference 99 /note= "encoded by CAA"
 FT Misc-difference 103 /note= "encoded by CCC"
 FT Misc-difference 141 /note= "encoded by GGC"
 FT Misc-difference 159 /note= "encoded by AAC"
 FT Misc-difference 171 /note= "encoded by ATG"
 FT Misc-difference 176 /note= "encoded by GTC"
 FT Misc-difference 256 /note= "encoded by GGC"
 FT Misc-difference 256 /note= "encoded by GGC"

XX US5830731-A.

XX 03-NOV-1998.

XX 21-MAY-1997; 97US-0861205.

XX 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

XX 13-JUL-1989; 89US-0379076.

XX 13-JUL-1990; 90US-0553759.

PR 21-MAY-1997; 97US-0861205.

XX (GENO) GEN HOSPITAL CORP.

XX Aruffo A, Seed B;

XX WPI: 1998-609251/51.

XX N-PSDB; V63456.

XX New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian CDNA(s), especially
 PT human lymphocyte antigenic sequences

PS Example 10; Column 53-54; 75pp; English.

XX This is the amino acid sequence of human Fc receptor I (FCRI), as
 CC deduced from cDNA clone p135 (see V63456). Isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from polyA RNA of cells
 CC from a single patient undergoing extracorporeal interleukin-2
 CC induction therapy. Clones p90 (see X07372) and p98 (see X07373)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-53). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.

SO Sequence 374 AA;

Query Match 75.0%; Score 48; DB 19; Length 374;
 Best Local Similarity 83.3%; Pred. No. 3.8;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDSGLYQCEAT 12
 DB 253 edsglywceat 264

RESULT 14

W97833 ID W97833 standard; Protein: 374 AA.

AC W97833;

DT 07-JUN-1999 (first entry)

DE Human Fc receptor I.

XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
 XX cloning.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 2 /note= "encoded by TGC"
 FT Misc-difference 23 /note= "encoded by CTC"
 FT Misc-difference 44 /note= "encoded by GAC"
 FT Misc-difference 45 /note= "encoded by GAC"

PT New cloning vector and polylinker - based on existing sequences for
PT efficient cloning and expression of mammalian cDNA(s), especially
PT human lymphocyte antigenic sequences

PS Example 10; Column 53-54; 75pp; English.

XX This is the amino acid sequence of human Fc receptor I (FcR1), as
CC deduced from cDNA clone p98/X2 (see X07373) isolated using a rapid
CC immunoselection cloning method from a cDNA library expressed in COS
CC cells. The cDNA library was constructed from polyA RNA of cells
CC from a single patient undergoing extracorporeal interleukin-2
CC induction therapy. Clones p135 (see V63456) and p90 (see X07372)
CC were also obtained. A novel method for cloning cDNAs from mammalian
CC expression libraries is based on transient expression of an antigen
CC in eukaryotic cells and selection of cells expressing the antigen by
CC adhesion to an antibody-coated substrate. The method is useful for
CC the isolation and cloning of any protein which can be expressed and
CC transported to the cell surface membrane of a eukaryotic cell. It
CC has been used to clone genes (see V63442-63) encoding cell surface
CC antigens from mammalian lymphocytes (see W80440-55). The isolated
CC genes can be expressed in a prokaryotic or eukaryotic host cells to
CC produce the encoded protein. The invention also provides high
CC efficiency expression vectors (see V63441 and V63444) which allow
CC the generation of very large mammalian expression libraries. The
CC purified genes and proteins are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders of
CC animals, including humans.

XX
SQ Sequence 374 AA;

Query Match 75.0%; Score 48; DB 19; Length 374;
Best Local Similarity 83.3%; Pred. No. 3.8;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDSGLYCEAAT 12
||| |||||
Db 253 edsclwyceaat 264

Search completed: June 4, 2001, 12:13:31
Job time: 204 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:31 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-26

Perfect score: 64

Sequence: 1 EDSCLYOCCEAAT 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6C_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	75.0	605	2	US-08-752-307B-8
2	48	75.0	1018	1	US-08-452-052-2
3	47	73.4	1447	5	PCT-US94-05277-2
4	46	71.9	75	4	US-08-928-383B-13
5	46	71.9	91	2	US-08-332-562A-94
6	46	71.9	607	2	US-08-752-307B-12
7	46	71.9	1018	1	US-08-408-093-6
8	46	71.9	1018	1	US-08-408-420A-6
9	46	71.9	1018	1	US-08-714-901-6
10	46	71.9	1018	3	US-08-040-741-6
11	45	70.3	72	4	US-08-928-383B-11
12	43	67.2	272	1	US-08-882-951-6
13	43	67.2	767	2	US-08-874-678-2
14	43	67.2	767	2	US-08-643-839-2
15	43	67.2	828	1	US-08-261-304-2
16	43	67.2	1356	1	US-08-810-116-8
17	43	67.2	1356	2	US-07-930-548A-8
18	43	67.2	1356	4	US-09-098-707A-2
19	42	65.6	643	5	PCT-US93-00031-19
20	42	65.6	644	5	PCT-US93-00031-21
21	42	65.6	647	5	PCT-US93-00031-11
22	42	65.6	647	5	PCT-US93-00031-23
23	42	65.6	662	1	US-08-261-304-7
24	42	65.6	735	5	PCT-US93-00031-13
25	42	65.6	736	5	PCT-US93-00031-15
26	42	65.6	739	5	PCT-US93-00031-9
27	42	65.6	805	3	US-08-985-526-34

28	42	65.6	806	2	US-08-443-861-5	Sequence 5, Appl1
29	42	65.6	806	4	US-08-193-829B-5	Sequence 5, Appl1
30	42	65.6	1367	1	US-07-813-593-4	Sequence 4, Appl1
31	42	65.6	1367	1	US-07-977-451-6	Sequence 6, Appl1
32	42	65.6	1367	1	US-07-946-507-4	Sequence 4, Appl1
33	42	65.6	1367	1	US-08-252-517-6	Sequence 6, Appl1
34	42	65.6	1367	1	US-07-906-397A-6	Sequence 6, Appl1
35	42	65.6	1367	1	US-08-601-891-6	Sequence 6, Appl1
36	42	65.6	1367	2	US-08-443-861-2	Sequence 2, Appl1
37	42	65.6	1367	2	US-09-021-324-6	Sequence 6, Appl1
38	42	65.6	1367	4	US-08-193-829B-2	Sequence 2, Appl1
39	42	65.6	1367	5	PCT-US92-02750-8	Sequence 8, Appl1
40	42	65.6	1367	5	PCT-US92-05401-6	Sequence 6, Appl1
41	42	65.6	1367	5	PCT-US92-09893-6	Sequence 6, Appl1
42	42	64.1	90	4	US-08-928-383B-22	Sequence 22, Appl1
43	41	64.1	119	1	US-08-497-025-6	Sequence 6, Appl1
44	41	64.1	740	5	PCT-US93-00031-17	Sequence 17, Appl1
45	41	64.1	1021	1	US-08-497-025-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-752-307B-8
Sequence 8, Application US/08752307B
Patent No. 5952171
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Geating, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
TITLE OF INVENTION: ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,307B
FILING DATE: 19-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 09404/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-752-307B-8
Query Match 75.0%; Score 48; DB 2; Length 605;
Best local Similarity 80.0%; Pred. No. 2.1;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDSGLYCCEA 10
11111111
Db 303 EDEGLYCCEA 312

RESULT 2
US-08-452-052-2
Sequence 2, Application US/08452052
Patent No. 5766922

GENERAL INFORMATION:

APPLICANT: Peles, Elior
TITLE OF INVENTION: FUNCTIONAL LIGANDS FOR THE AXONAL CELL
TITLE OF INVENTION: RECOGNITION MOLECULE CONTRACTIN
NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/452,052

FILING DATE: 26-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Cortuzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7683-075

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1018 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-452-052-2

Query Match 75.0%; Score 48; DB 1; Length 1018;
Best Local Similarity 80.0%; Pred. No. 3.6;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDSGLYCCEA 10
11111111
Db 302 EDEGLYCCEA 311

RESULT 3
PCT-US94-05277-2
Sequence 2, Application PC/TUS9405277
GENERAL INFORMATION:

APPLICANT: Bruskin, Arthur

APPLICANT: Jarosz, David E.

APPLICANT: Johnson, Karen

APPLICANT: Kinzler, Kenneth W.

APPLICANT: Vogelstein, Bert

APPLICANT: Zablacky, James R.

TITLE OF INVENTION: Antibodies Specific for DCC Gene Product

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner, Birch, McKie & Beckett

STREET: 1001 G Street, N.W.

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/05277

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 01107.42709

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202.508.9100

TELEFAX: 202.508.9299

TELEX: 197430 BBMB UT

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1447 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-05277-2

Query Match 73.4%; Score 47; DB 5; Length 1447;
Best Local Similarity 80.0%; Pred. No. 7.5;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 DSGLYCCEA 11
11111111
Db 111 DGLYCCEAS 120

RESULT 4
US-08-928-383B-13

Sequence 13, Application US/08928383B

Patent No. 6210921

GENERAL INFORMATION:

APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,

APPLICANT: and Marshall S. Horwitz.

TITLE OF INVENTION: CAR, A No. 6210921el Coxsackievirus and Adenovirus

TITLE OF INVENTION: Receptor

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,383B

FILING DATE: 12-SEP-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/026,100

FILING DATE: 13-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: DFN-020

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ. ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 75 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-928-383B-13

Query Match 71.9%; Score 46; DB 4; Length 75;
Best Local Similarity 70.0%; Pred. No. 0.51;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDSGLYCCEA 10
DB 48 EDSGLYCCEA 57

RESULT 5
US-08-332-562A-94
Sequence 94, Application US/08332562A
Patent No. 5985599
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HUETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ. ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-94

Query Match 71.9%; Score 46; DB 2; Length 91;
Best Local Similarity 66.7%; Pred. No. 0.63;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDSGLYCCEAT 12
DB 65 EDSGLYCCEAT 76

RESULT 6
US-08-752-307B-12
Sequence 12, Application US/08752307B
Patent No. 5952171

GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearling, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
TITLE OF INVENTION: ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,307B
FILING DATE: 19-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkijohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 09404/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-752-307B-12

Query Match 71.9%; Score 46; DB 2; Length 607;
Best Local Similarity 80.0%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDSGLYCCEA 10
DB 388 EDSGLYCCEA 397

RESULT 7
US-08-408-093-6
Sequence 6, Application US/08408093
Patent No. 5688916
GENERAL INFORMATION:
APPLICANT: Reid, Robert A.
APPLICANT: Hemperly, John J.
TITLE OF INVENTION: Human Cell Adhesion Molecule and Nucleic
TITLE OF INVENTION: Acid Sequences
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSEE: Richard J. Rodrick, Becton Dickinson
 ADDRESSEE: and Company
 STREET: One Becton Drive
 CITY: Franklin Lakes
 STATE: NJ
 COUNTRY: USA
 ZIP: 07417

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/408,093
 FILING DATE: 21-MAR-1995
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/040,741
 FILING DATE: 26 MAR 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Fugit, Donna R.
 REGISTRATION NUMBER: 32,135
 REFERENCE/DOCKET NUMBER: P-2630
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1018 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 45..94
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 138..191
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 243..290
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 332..371
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 416..464
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 506..563
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 604..657
 OTHER INFORMATION: /note= "conserved core of fibronectin type
 OTHER INFORMATION: IIT-like repeat"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 707..760
 OTHER INFORMATION: /label= FLR
 OTHER INFORMATION: /note= "conserved core of fibronectin type
 OTHER INFORMATION: IIT-like repeat"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 809..857
 OTHER INFORMATION: /label= FLR
 OTHER INFORMATION: /note= "conserved core of fibronectin type
 OTHER INFORMATION: IIT-like repeat"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 905..952
 OTHER INFORMATION: /label= FLR
 OTHER INFORMATION: /note= "conserved core of fibronectin type
 OTHER INFORMATION: IIT-like repeat"

FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 188
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 238
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 318
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 437
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 453
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 474
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 501
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 571
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 913
 OTHER INFORMATION: /label= ASN-glycos
 OTHER INFORMATION: /note= "potential site of ASN-linked
 glycosylation"
 US-08-408-093-6

Query Match 71.9% Score 46; DB 1; Length 1018;
 Best Local Similarity 70.0% Pred. NO. 7.6;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EDGGLYQCEA 10
 |||:|:|
 Db 303 EDEGLYQCEA 312

RESULT 8
 US-08-408-420A-6
 : Sequence 6, Application US/08408420A
 : Patent No. 5731154
 : GENERAL INFORMATION:
 : APPLICANT: Reid, Robert A.
 : APPLICANT: Hemperly, John J.

TITLE OF INVENTION: Human Cell Adhesion Molecule and Nucleic
TITLE OF INVENTION: Acid Sequences
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
ADDRESSEE: Company
STREET: One Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,420A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-2630
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1018 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 45..94
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 138..191
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 243..290
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 332..371
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 416..464
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 506..563
FEATURE:
NAME/KEY: Domain
LOCATION: 604..657
OTHER INFORMATION: /label- FLR
OTHER INFORMATION: /note- "conserved core of fibronectin type
OTHER INFORMATION: IIT-like repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 707..760
OTHER INFORMATION: /label- FLR
OTHER INFORMATION: /note- "conserved core of fibronectin type
OTHER INFORMATION: IIT-like repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 809..857
OTHER INFORMATION: /label- FLR
OTHER INFORMATION: /note- "conserved core of fibronectin type
OTHER INFORMATION: IIT-like repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 905..952
OTHER INFORMATION: /label- FLR
OTHER INFORMATION: /note- "conserved core of fibronectin type

OTHER INFORMATION: IIT-like repeat"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 188
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 238
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 318
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 437
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 453
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 474
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 501
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 571
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 913
OTHER INFORMATION: /label- ASN-glycos
OTHER INFORMATION: /note- "potential site of ASN-linked
OTHER INFORMATION: glycosylation"
US-08-408-420A-6
Query Match 71.9%; Score 46; DB 1; Length 1018;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 EDGGLYQCEA 10
Db 303 EDEGIYCEA 312
RESULT 9
US-08-714-901-6
Sequence 6, Application US/08714901
Patent No. 5739289
GENERAL INFORMATION:
APPLICANT: Reid, Robert A.

APPLICANT: Hemperly, John J.
 TITLE OF INVENTION: Human Cell Adhesion Molecule and Nucleic
 NUMBER OF INVENTION: Acid sequences
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Richard J. Rodrick, Becton Dickinson
 STREET: One Becton Drive
 CITY: Franklin Lakes
 STATE: NJ
 COUNTRY: USA
 ZIP: 07417
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/714,901
 FILING DATE: 17-SEP-1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/408,093
 FILING DATE: 21-MAR-1995
 APPLICATION NUMBER: US/08/040,741
 FILING DATE: 26 MAR 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Fugitt, Donna R.
 REGISTRATION NUMBER: 32,135
 REFERENCE/DOCKET NUMBER: P-2630
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1018 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 45..94
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 138..191
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 243..290
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 332..371
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 416..464
 FEATURE:
 NAME/KEY: Disulfide-bond
 LOCATION: 506..563
 NAME/KEY: Domain
 LOCATION: 707..760
 OTHER INFORMATION: /label=FLR
 OTHER INFORMATION: /note="conserved core of fibronectin type
 OTHER INFORMATION: IIT-like repeat"
 NAME/KEY: Domain
 LOCATION: 809..857
 OTHER INFORMATION: /label=FLR
 OTHER INFORMATION: /note="conserved core of fibronectin type

OTHER INFORMATION: IIT-like repeat"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 905..952
 OTHER INFORMATION: /label=FLR
 OTHER INFORMATION: /note="conserved core of fibronectin type
 OTHER INFORMATION: IIT-like repeat"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 188
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 238
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 318
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 437
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 453
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 474
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 501
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 571
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 913
 OTHER INFORMATION: /label=ASN-glycos
 OTHER INFORMATION: /note="potential site of ASN-linked
 OTHER INFORMATION: glycosylation"
 US-08-714-901-6

Query Match 71.9%; Score 46; DB 1; Length 1018;
 Best Local Similarity 70.0%; Pred. No. 7.6;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDSGIGCEA 10
 DB 303 EDSGIGCEA 312

RESULT 10
US-08-040-741-6
Sequence 6, Application US/08040741
Patent No. 6017695
GENERAL INFORMATION:
APPLICANT: Reid, Robert A.
APPLICANT: Hemperly, John J.
TITLE OF INVENTION: Human Cell Adhesion Molecule and Nucleic
Acid Sequences
TITLE OF INVENTION: Acid Sequences
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard J. Rodrick, Becton Dickinson
ADDRESSEE: and Company
STREET: One Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: USA
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/040,741
FILING DATE: 19930326
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Fugitt, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-2630
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1018 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 45..94
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 138..191
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 243..290
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 332..371
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 416..464
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 506..563
FEATURE:
NAME/KEY: Domain
LOCATION: 604..657
OTHER INFORMATION: /label= FLR
OTHER INFORMATION: /note= "conserved core of fibronectin type
III-like repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 707..760
OTHER INFORMATION: /label= FLR
OTHER INFORMATION: /note= "conserved core of fibronectin type
III-like repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 809..857
OTHER INFORMATION: /label= FLR

OTHER INFORMATION: /note= "conserved core of fibronectin type
III-like repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 905..952
OTHER INFORMATION: /label= FLR
OTHER INFORMATION: /note= "conserved core of fibronectin type
III-like repeat"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 188
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 238
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 318
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 437
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 453
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 474
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 501
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 571
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 913
OTHER INFORMATION: /label= ASN-glycos
OTHER INFORMATION: /note= "potential site of ASN-linked
glycosylation"
OTHER INFORMATION: glycosylation"
US-08-040-741-6

Query Match 71.9%; Score 46; DB 3; Length 1018;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 EDSGLYQCEA 10
Db 303 EDECIYCEA 312

RESULT 11
US-08-928-383B-11
Sequence 11, Application US/08928383B
Patent No. 6210921
GENERAL INFORMATION:
APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,
APPLICANT: and Marshall S. Horwitz
TITLE OF INVENTION: CAR, A No. 6210921el Cocksackievirus and Adenovirus
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,383B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,100
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DEN-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)742-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-928-383B-11

Query Match 70.3%; Score 45; DB 4; Length 72;
Best Local Similarity 50.0%; Pred: No. 0.72;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSLYQCEAAT 12
DB 46 ODAGIVECESKT 57

RESULT 12
US-08-282-951-6
Sequence 6, Application US/08282951
Patent No. 5665590
GENERAL INFORMATION:
APPLICANT: YANG, ZHI
TITLE OF INVENTION: METHOD FOR ISOLATING AND DIRECTLY
TITLE OF INVENTION: CLONING GENES WHICH ENCODE CELL-SURFACE AND SECRETED
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/282,951
FILING DATE: 29-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 20296-20012.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 272 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-282-951-6

Query Match 67.2%; Score 43; DB 1; Length 272;
Best Local Similarity 60.0%; Pred: No. 5.9;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSLYQCEA 10
DB 86 EDAGIRYRCA 95

RESULT 13
US-08-874-678-2
Sequence 2, Application US/08874678
Patent No. 5952199
GENERAL INFORMATION:
APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Presta, Leonard
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR
TITLE OF INVENTION: ACTIVITY, THEIR USES AND PROCESSES FOR THEIR PRODUCTION
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albrighton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,678
FILING DATE: HEREWITH
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/643,839
FILING DATE: 07-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63291-1/MHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249

TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 767 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-874-678-2

Query Match 67.2%; Score 43; DB 2; Length 767;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGLGYOCEAT 12
Db 730 EDEGLYCOACS 741

RESULT 14
US-08-643-839-2
Sequence 2, Application US/08643839
Patent No. 6100071
GENERAL INFORMATION:
APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Presta, Leonard
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
TITLE OF INVENTION: PRODUCTION
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Honbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/643,839
FILING DATE: 07-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63291/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 767 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-643-839-2

Query Match 67.2%; Score 43; DB 3; Length 767;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGLGYOCEAT 12
|| ||| ||| :

Db 730 EDEGLYCOACS 741

RESULT 15
US-08-261-304-2
Sequence 2, Application US/08261304
Patent No. 5708147
GENERAL INFORMATION:
APPLICANT: Cybulsky, Myron I.
APPLICANT: Gimbrone, Michael A.
APPLICANT: Collins, Tucker
TITLE OF INVENTION: Mononuclear Leukocyte Directed
TITLE OF INVENTION: Endothelial Adhesion Molecule Associated with
TITLE OF INVENTION: Atherosclerosis
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W.
STREET: Suite 300
CITY: Washington
STATE: District of Columbia
COUNTRY: United States of America
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Ascii
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,304
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/649,565
FILING DATE: 01-FEB-1991
APPLICATION NUMBER: U.S. 07/487,038
FILING DATE: 02-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 0627, 2100004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 833-7533
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 828 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-261-304-2

Query Match 67.2%; Score 43; DB 1; Length 828;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGLGYOCEA 10
|| ||| ||| :
Db 661 EDAGVYECES 670

Search completed: June 4, 2001, 12:14:31
Job time: 224 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:43 ; Search time 65.22 seconds

(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-26

Perfect score: 64

Sequence: 1 EDGSLYQCEAAT 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935.residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

PIR 67: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	DB	ID	Description
1	57	89.1	344	2	A41357	Fc gamma (Igg) rec
2	57	89.1	374	1	A39878	Fc gamma (Igg) rec
3	49	76.6	1232	2	T43027	neural cell adhesi
4	48	75.0	555	2	J01526	interleukin-1 rece
5	48	75.0	1020	2	S05944	neutonal cell surf
6	48	75.0	1021	2	A57112	contactin precursor
7	48	75.0	1091	2	S01998	contactin precursor
8	48	75.0	1256	2	T03096	CDO protein - rat
9	47	73.4	267	2	A38442	probable tumor sup
10	47	73.4	416	2	A54017	colon carcinoma-as
11	47	73.4	1287	2	T30988	hypothetical prote
12	47	73.4	1447	2	A54100	tumor suppressor p
13	46	71.9	336	2	I48471	Fc gamma (Igg) rec
14	46	71.9	404	2	A46480	Fc gamma (Igg) rec
15	46	71.9	1018	2	JC4211	neural adhesion pr
16	46	71.9	1018	2	A54744	contactin 1 precur
17	46	71.9	1036	2	S22383	axonin 1 precursor
18	46	71.9	1040	2	A34695	axonal glycoprotei
19	46	71.9	1040	2	A49356	transient axonal g
20	45	70.3	475	2	I76668	pregnancy-specific
21	45	70.3	739	2	JN0581	vascular cell adhe
22	45	70.3	1197	2	T30581	neural cell adhesi
23	44	68.8	162	2	I51668	tumor suppressor
24	44	68.8	739	2	J50675	vascular cell adhe
25	44	68.8	1427	2	I51669	tumor suppressor
26	44	68.8	7962	2	I38346	elastic titin - hu
27	43	67.2	137	2	E34903	ig heavy chain pre
28	43	67.2	194	2	T29925	hypothetical prote
29	43	67.2	282	2	T17219	hypothetical prote

30	43	67.2	500	2	S59795	hypothetical prote
31	43	67.2	518	2	JC4024	poliovirus recepto
32	43	67.2	1015	2	T32186	hypothetical prote
33	43	67.2	1356	2	JC1402	protein-tyrosine k
34	43	67.2	26926	1	I38344	titin, cardiac mus
35	42	65.6	475	2	A54879	pregnancy-specific
36	42	65.6	538	2	JC2457	vascular cell adhe
37	42	65.6	647	2	B41288	vascular cell adhe
38	42	65.6	739	2	A41288	protein-tyrosine k
39	42	65.6	790	2	A39627	fibroblast growth
40	42	65.6	820	2	S17295	fibroblast growth
41	42	65.6	821	1	TVMSBK	fibroblast growth
42	42	65.6	822	2	B54846	cell adhesion prot
43	42	65.6	1033	2	S19247	protein-tyrosine k
44	42	65.6	1052	2	B49120	hypothetical prote
45	42	65.6	1173	2	T25893	

ALIGNMENTS

```

RESULT 1
A41357
Fc gamma (Igg) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A:title: Isolation and expression of functional high-affinity Fc receptor complementa
A:Reference number: A41357; MID:89100284
A:Accession: A41357
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: GB:X14355; GB:M21090; MID:931333; PIDN:CAA32536.1; PID:931334
R:Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A:title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (
A:Reference number: S03018; MID:89098339
A:Accession: S03019
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: EMBL:X14355; MID:931333; PIDN:CAA32536.1; PID:931334
A:Note: the authors translated the codon ACT for residue 25 as Ala
C:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
F:117-170/Domain: immunoglobulin homology <IM>

Query Match 89.1%; Score 57; DB 2; Length 344;
Best local Similarity 91.7%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSLYQCEAAT 12
DB 253 EDGSLYQCEAAT 264

RESULT 2
A39878
Fc gamma (Igg) receptor I-A (high affinity) precursor - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: A39878; MID:91302383
R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
J. Biol. Chem. 266, 13449-13455, 1991
A:title: Gene organization of the human high affinity receptor for IgG, Fc gamma 1 (CD
A:Reference number: A39878; MID:91302383
A:Accession: A39878
A:Molecule type: DNA
A:Residues: 1-374 <VAN>

```

A:Cross-references: GB:M63830; GB:M63835; NID:g180227; PIDN:AAA5678.1; PID:g180229
 R:Portes, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: 155577; MUID:93055454
 A:Accession: 170304
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:L03418; NID:g184840; PIDN:AAA36049.1; PID:g292169
 R:Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: M41357; MUID:89100284
 A:Accession: B41357
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'T', '339-374' <AL1>
 A:Cross-references: GB:X14356; GB:M21091; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R:Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR)
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'T', '339-374' <AL2>
 A:Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R:Perez, C.; Wietzeblin, J.; Benech, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A:Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
 ism.
 A:Reference number: 157525; MUID:93204964
 A:Accession: 157525
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RE2>
 A:Cross-references: GB:S57204; NID:g298692; PIDN:AAD13887.1; PID:g4261587
 C:Genetics:
 A:Gene: GDB:FCGR1A; CDB4
 A:Cross-references: GDB:135911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane F
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-292/Domain: extracellular #status predicted <EXT>
 F:117-170/Domain: immunoglobulin homology <IMM2>
 F:293-313/Domain: transmembrane #status predicted <TM>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.1%; Score 57; DB 1; Length 374;
 Best Local Similarity 91.7%; Pred. No. 0 018;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 253 EDSGLVCEAAT 264

Y 1 EDSGLVCEAAT 12
 ||||| |||||
 ||||| |||||

RESULT 3
 T43027
 neural cell adhesion molecule L1 - goldfish
 N:Alternate names: E587 antigen
 C:Species: Carassius auratus (goldfish)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T43027
 R:Giordano, S.; Laessing, U.; Lottspeich, F.; Stuermer, C.A.O.
 submitted to the EMBL Data Library, April 1996
 A:Description: Molecular cloning of goldfish E587 antigen, a cell adhesion molecule exp
 A:Reference number: 222294
 A:Accession: T43027
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1232 <GIO>

A:Cross-references: EMBL:U55211; NID:g1305526; PID:g1305527; PIDN:AAA99159.1
 C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology
 C:Keywords: cell adhesion; membrane protein

Query Match 76.6%; Score 49; DB 2; Length 1232;
 Best Local Similarity 66.7%; Pred. No. 1.3;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Y 1 EDSGLVCEAAT 12
 ||||| |||||
 ||||| |||||

Db 553 EDSGLVCEAAT 564

Y 1 EDSGLVCEAAT 12
 ||||| |||||
 ||||| |||||

RESULT 4
 J01526
 Interleukin-1 receptor I precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 29-Sep-1999
 C:Accession: J01526
 R:Gulida, S.; Heguy, A.; Mell, M.
 Gene 111, 229-243, 1992
 A:Title: The chicken IL-1 receptor: differential evolution of the cytoplasmic and ext
 A:Reference number: J01526; MUID:92175529
 A:Accession: J01526
 A:Molecule type: mRNA
 A:Residues: 1-555 <GU>
 A:Cross-references: GB:M81846; NID:g212206; PIDN:AAA8924.1; PID:g212207
 A:Experimental source: embryo
 C:Superfamily: interleukin-1 receptor type I
 C:Keywords: cytokine receptor; transmembrane protein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-555/Product: interleukin-1 receptor I #status predicted <INT>
 F:339-359/Domain: transmembrane #status predicted <TRA>

Query Match 75.0%; Score 48; DB 2; Length 555;
 Best Local Similarity 88.9%; Pred. No. 0.93;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 EDSGLVCEA 9
 ||||| |||||
 ||||| |||||

Db 91 EDSGLVCEA 99

Y 1 EDSGLVCEA 9
 ||||| |||||
 ||||| |||||

RESULT 5
 S05944
 neuronal cell surface protein F3 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 21-Jan-2000
 C:Accession: S05944
 R:Gennari, G.; Cibelli, G.; Rougon, G.; Mattei, M.C.; Goridis, C.
 J. Cell Biol. 109, 775-788, 1989
 A:Title: The mouse neuronal cell surface protein F3: a phosphatidylinositol-anchored
 A:Reference number: S05944; MUID:89340657
 A:Accession: S05944
 A:Molecule type: mRNA
 A:Residues: 1-1020 <GEN>
 A:Cross-references: EMBL:X14943; NID:g50937; PIDN:CAA33075.1; PID:g50938
 C:Genetics:
 A:Map position: 15F
 C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homolo
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:256-312/Domain: immunoglobulin homology <IMM>

Query Match 75.0%; Score 48; DB 2; Length 1020;
 Best Local Similarity 80.0%; Pred. No. 1.6;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 EDSGLVCEA 10
 ||||| |||||
 ||||| |||||

Db 303 EDSGLVCEA 312


```

RESULT 6
A:57112
C:contactin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C:Accession: A57112
R:Peles, E.; Nativ, M.; Campbell, P.L.; Sakurai, T.; Martinez, R.; Lev, S.; Clary, D.O.;
Cell 82, 251-260, 1995
A:Title: The carbonic anhydrase domain of receptor tyrosine phosphatase beta is a functionally
A:Reference number: A57112; PMID:95354206
A:Accession: A57112
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-1021 <PEL>
C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology
C:Keywords: membrane protein; phosphatidylinositol linkage
F:256-312/Domain: immunoglobulin homology <IMM>

Query Match 75.0%; Score 48; DB 2; Length 1021;
Best Local Similarity 80.0%; Pred. No. 1.6;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSLYCEA 10
DB 303 EDGSLYCEA 312

RESULT 7
S01998
C:contactin precursor - chicken
N:Alternate names: 130K glycoprotein
C:Species: Gallus gallus (chicken)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 17-Nov-2000
C:Accession: S01998; J00094
R:Ranscht, B.; Dours, M.T.;
J. Cell Biol. 107, 1561-1573, 1988
A:Title: Sequence of contactin, a 130-KD glycoprotein concentrated in areas of interneur
A:Reference number: S01998; PMID:89008597
A:Accession: S01998
A:Molecule type: mRNA
A:Residues: 1-1091 <RAN>
A:Cross-references: EMBL:Y00813; NID:g63328; PIDN:CA68753.1; PID:g63329
A:Note: part of this sequence, including the amino end of the mature protein, was confir
R:Brudemendorf, T.; Wolff, J.M.; Frank, R.; Rathjen, F.G.
Neuron 2, 1351-1361, 1989
A:Title: Neutral cell recognition molecule F11: homology with fibronectin type III and in
A:Reference number: J00094; PMID:90180453
A:Accession: J00094
A:Molecule type: mRNA
A:Residues: 1-970, 'S', 972-1000, 1090-1091, 'GVLAYSGF' <BRU>
A:Cross-references: GB:X14877; NID:91708784; PIDN:CA33018.1; PID:g63385
A:Note: the carboxy-end hydrophobic stretch is compatible with the consensus motif for c
A:Note: F11 comprises six domains related to the immunoglobulin domain type C and four i
CAM
C:Comment: F11 is a chick neural cell surface-associated glycoprotein implicated in neur
C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology
C:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linkag
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-1091/Product: contactin #status predicted <MNT>
F:21-982/Domain: extracellular #status predicted <EXT>
F:247-303/Domain: immunoglobulin homology <IMM>
F:983-1003/Domain: transmembrane #status predicted <TM>
F:1003-1091/Domain: intracellular #status predicted <INT>
F:200, 249, 329, 448, 485, 512, 582, 621, 924/binding site: carbohydrate (Asn) (covalent) #S
F:984/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 75.0%; Score 48; DB 2; Length 1091;
Best Local Similarity 80.0%; Pred. No. 1.7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 EDGSLYCEA 10
DB 294 EDGSLYCEA 303

RESULT 8
T03096
C:CDO protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 05-Nov-1999
C:Accession: T03096
R:Krauss, R.S.; Kang, J.S.; Gao, M.; Feinleib, J.L.;
submitted to the EMBL data library, September 1998
A:Reference number: Z14837
A:Accession: T03096
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1256 <KRA>
A:Cross-references: EMBL:AF004840; NID:g3550957; PIDN:AC34735.1; PID:g3550958
A:Experimental source: fibroblast
C:Genetics:
A:Note: cdo

Query Match 75.0%; Score 48; DB 2; Length 1256;
Best Local Similarity 90.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSLYCEA 10
DB 373 EDGSLYCEA 382

RESULT 9
A38442
C:probable tumor suppressor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 16-Feb-1997
C:Accession: A38442
R:Nigro, J.M.; Cho, K.R.; Fearon, E.R.; Kern, S.E.; Ruppert, J.M.; Oliner, J.D.; Kinz
Cell 64, 607-613, 1991
A:Title: Scrambled exons.
A:Reference number: A38442; PMID:91121517
A:Accession: A38442
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <NIG>
A:Cross-references: GB:M63696; GB:M63700; GB:M63702; GB:M63718; GB:M63698
C:Keywords: transmembrane protein

Query Match 73.4%; Score 47; DB 2; Length 267;
Best Local Similarity 80.0%; Pred. No. 0.7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 DSGLYCEA 11
DB 5 DSGLYCEA 14

RESULT 10
A54017
C:colon carcinoma-associated antigen pE4 precursor - rat
N:Alternate names: pE4 protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 01-Dec-1995
C:Accession: A54017; A61206
R:Chadenau, C.; Lemoullac, B.; Denis, M.G.
J. Biol. Chem. 269, 15601-15605, 1994
A:Title: A novel member of the immunoglobulin gene superfamily expressed in rat carc
A:Reference number: A54017; PMID:94253144
A:Accession: A54017

```

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-416 <CHN>
 A:Cross-references: GB:LI2025
 R:Chadenet, C.; Denis, M.G.; Blotiere, H.M.; Gregoire, M.; Douillard, J.Y.; Meflah, K.
 Int. J. Cancer 47, 903-908, 1991
 A:Title: Characterization, isolation and amino terminal sequencing of a rat colon carcin
 A:Reference number: A61206; MUID:91184910
 A:Accession: A61206
 A:Molecule type: protein
 A:Residues: 34-41, 'X', 43-53 <CH2>
 A:Note: the residue at position 9 is suggested to be glycosylated asparagine
 C:Comment: This 66k protein is detected by monoclonal antibody E4 on colon carcinoma cell
 C:Keywords: glycoprotein; membrane protein

Query Match 73.4%; Score 47; DB 2; Length 416;
 Best Local Similarity 58.3%; Pred. No. 1.1;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDSGLYCEAAT 12
 |||:|:|:|
 Db 122 EDEGIYECQIAT 133

RESULT 11

hypothetical protein C18F3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000

C:Accession: T30988

R:Du, Z.; Gattung, S.

submitted to the EMBL Data Library, August 1999

A:Description: The sequence of C. elegans cosmid C18F3.

A:Reference number: Z20953

A:Accession: T30988

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1287 <DNZ>

A:Cross-references: EMBL:U50067; PIDN:AAA93439.2

A:Experimental source: strain Bristol N2

C:Genetics:

A:Map position: IV

A:Introns: 34/3; 72/1; 131/2; 157/1; 254/3; 303/1; 400/1; 495/1; 592/1; 681/3; 950/1; 10

A:Note: C18F3.2

C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; I

Query Match 73.4%; Score 47; DB 2; Length 1287;
 Best Local Similarity 88.9%; Pred. No. 3;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DSGLYQCEA 10
 ||:|||||
 Db 168 DAGLYQCEA 176

RESULT 12

tumor suppressor protein DCC precursor - human

N:Alternate names: colorectal cancer suppressor DCC

C:Species: Homo sapiens (man)

C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 05-Nov-1999

C:Accession: A54100; A40098

R:Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.

Genes Dev. 8, 1174-1183, 1994

A:Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.

A:Reference number: A54100; MUID:95011532

A:Accession: A54100

A:Molecule type: mRNA

A:Residues: 1-1447 <HED>

A:Cross-references: EMBL:X76132; NID:9453209; PIDN:CAA53735.1; PID:9453210

R:Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilton

Science 247, 49-56, 1990

A:Title: Identification of a chromosome 18q gene that is altered in colorectal cancer

A:Reference number: A40098; MUID:90100559

A:Accession: A40098

A:Molecule type: mRNA

A:Residues: 1-750 <FEA>

A:Cross-references: GB:M32292; NID:9181492; PIDN:AAA5751.1; PID:9181493

C:Genetics:

A:Gene: GDB:DCC

A:Cross-references: GDB:119838; OMIM:120470

A:Map position: 18q21.1-18q21.1

C:Keywords: transmembrane protein; tumor suppressor

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>

Query Match 73.4%; Score 47; DB 2; Length 1447;
 Best Local Similarity 80.0%; Pred. No. 3.4;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 DSGLYQCEA 11
 |||:|:|:|
 Db 111 DEGLYQCEAS 120

RESULT 13

Fc gamma (IgG) receptor high affinity - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000

C:Accession: I48471

R:Pins, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarth, P.M.; Wicker, L.S.; G

Science 260, 695-698, 1993

A:Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for

A:Reference number: I48471; MUID:93242599

A:Accession: I48471

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-336 <RES>

A:Cross-references: EMBL:X70980; NID:9311748; PIDN:CAA5031.1; PID:9311749

C:Superfamily: Fc gamma receptor I; Immunoglobulin homology

C:Keywords: Immunoglobulin receptor

F:128-180/Domain: Immunoglobulin homology <IMM>

Query Match 71.9%; Score 46; DB 2; Length 336;
 Best Local Similarity 66.7%; Pred. No. 1.3;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDSGLYCEAAT 12
 ||:|:|:|:|
 Db 267 EDAGFYWCQVAT 278

RESULT 14

A46480

Fc gamma (IgG) receptor high affinity - mouse

N:Alternate names: high affinity IgG receptor

C:Species: Mus musculus (house mouse)

C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000

C:Accession: A46480; A43511

R:Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.

J. Immunol. 148, 1570-1575, 1992

A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI a

A:Reference number: A46480; MUID:92166399

A:Accession: A46480

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-404 <OSM>

A:Note: Sequence extracted from NCBI backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212,

R:Seares, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.

J. Immunol. 144, 371-378, 1990

A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for

A:Reference number: A43511; MUID:90111035
A:Accession: A43511
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-404 <SEA>
A:Cross-references: GB:M31314; NID:g200752; PIDN:AAA40056.1; PID:g200753
C:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: immunoglobulin receptor; transmembrane protein
F:127-179/Domain: immunoglobulin homology <TM>

Query Match 71.9%; Score 46; DB 2; Length 404;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGSLYOCENAT 12
||:|:|:|:|:|
Db 262 EDAGFYWCCEVAT 273

RESULT 15

JC4211
neural adhesion protein F3 precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Jun-2000
C:Accession: JC4211
R:Matanabe, K.; Shinazaki, K.; Hosoya, H.; Fukumauchi, F.; Takenawa, T.
Gene 160, 245-248, 1995
A:title: Cloning of the cDNA encoding neural adhesion molecule F3 from bovine brain.
A:Reference number: JC4211; MUID:95369697
A:Accession: JC4211
A:Molecule type: mRNA
A:Residues: 1-1018 <MAT>
A:Cross-references: DDBJ:D32135; NID:g1060860; PIDN:BAA06861.1; PID:g1060861
A:Experimental source: brain
C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology
C:Keywords: brain; glycoprotein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-1018/Product: neural adhesion protein F3 #status predicted <MAT>
F:256-312/Domain: immunoglobulin homology <IMM>
F:992-1018/Region: hydrophobic #status predicted
F:208,258,338,457,473,494,521,591,630,933/Binding site: carbohydrate (Asn) (covalent) #

Query Match 71.9%; Score 46; DB 2; Length 1018;
Best Local Similarity 70.0%; Pred. No. 3.6;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSLYOCENAT 10
||:|:|:|:|:|
Db 303 EDGSLYECCEA 312

Search completed: June 4, 2001, 12:15:44
Job time: 277 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:06 ; Search time 37.56 Seconds
(without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-26

Perfect score: 64
Sequence: 1 EDSCGXOCEAAT 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	57	89.1	374	1 FCGL_HUMAN
2	48	75.0	1010	1 CONT_CHICK
3	48	75.0	1020	1 CONT_MOUSE
4	47	73.4	1447	1 DCC_HUMAN
5	47	73.4	1447	1 DCC_MOUSE
6	46	71.9	404	1 FCGL_MOUSE
7	46	71.9	1018	1 CONT_MOUSE
8	46	71.9	1036	1 AXOL_CHICK
9	46	71.9	1040	1 AXOL_HUMAN
10	46	71.9	1040	1 AXOL_MOUSE
11	44	68.8	739	1 VCAI_MOUSE
12	44	68.8	739	1 VCAI_RAT
13	44	68.8	837	1 NCML_MOUSE
14	43	67.2	837	1 NCML_HUMAN
15	43	67.2	1356	1 VGR2_MOUSE
16	42	65.6	739	1 VCAI_HUMAN
17	42	65.6	821	1 FGR2_MOUSE
18	42	65.6	1052	1 FGR2_MOUSE
19	42	65.6	1343	1 VGR2_MOUSE
20	42	65.6	1367	1 VGR2_MOUSE
21	42	65.6	1694	1 SM_MOUSE
22	42	65.6	1906	1 KML_MOUSE
23	41	64.1	233	1 GP42_MOUSE
24	41	64.1	521	1 GP42_MOUSE
25	41	64.1	530	1 PVR2_MOUSE
26	41	64.1	530	1 PVR2_MOUSE
27	41	64.1	1914	1 KML_HUMAN
28	40	62.5	338	1 LAMP_MOUSE
29	40	62.5	338	1 LAMP_MOUSE
30	40	62.5	338	1 LAMP_MOUSE
31	40	62.5	394	1 LAMP_MOUSE
32	40	62.5	417	1 PGB_MOUSE
33	40	62.5	519	1 PGB_MOUSE

34	40	62.5	819	1 FCGL_CHICK	P21804 gallus galli
35	40	62.5	912	1 PGB_BOVIN	O28062 bos taurus
36	40	62.5	2124	1 PGB_MOUSE	P07897 rattus norv
37	40	62.5	2132	1 PGB_MOUSE	O61282 mus musculu
38	40	62.5	2333	1 PGB_MOUSE	O28343 canis famill
39	39	60.9	139	1 KML_MOUSE	P79280 sus scrofa
40	39	60.9	777	1 SM_MOUSE	O95025 homo sapien
41	39	60.9	806	1 FCGL_MOUSE	P18460 gallus galli
42	39	60.9	821	1 FGR2_MOUSE	P21802 homo sapien
43	39	60.9	883	1 PGB_MOUSE	O61361 mus musculu
44	39	60.9	883	1 PGB_MOUSE	P55068 rattus norv
45	39	60.9	1136	1 TIE1_BOVIN	O06805 bos taurus

ALIGNMENTS

RESULT	1	FCGL_HUMAN	STANDARD	PRT	374 AA
ID	FCGL_HUMAN	P12314	P12315		
AC	01-OCT-1989 (Rel. 12, Created)				
DT	01-OCT-1989 (Rel. 12, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I) (CD64).				
GN	FCGRIA OR FCGR1 OR FCGL OR IGFR1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89098339; PubMed=2974947;				
RA	Allen J.M., Seed B.;				
RT	"Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcRI)."				
RL	Nucleic Acids Res. 16:11824-11824(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89100284; PubMed=2911749;				
RA	Allen J.M., Seed B.;				
RT	"Isolation and expression of functional high-affinity Fc receptor complementary DNAs."				
RL	Science 243:378-381(1989).				
CC	- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN GAMMA. HIGH AFFINITY RECEPTOR.				
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.				
CC	- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.				
CC	- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.				
CC	- DATABASE: NAME=PROV; NOTE=CD guide CD64 entry;				
CC	WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd64.htm"				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce/ or send an email to license@isb-sdb.ch).				
CC	-----				
DR	EMBL: X14356; CAA32537.1; -				
DR	EMBL: X14355; CAA32536.1; -				
DR	PIR: S03018; S03018				
DR	PIR: S03019; S03019				
DR	PIR: A41357; A41357				
DR	PIR: B41357; B41357				
DR	HSSP: P12319; 1ALT				
DR	MIM: 146760; -				
DR	InterPro: IPR003006; -				
DR	Pfam: PF00047; 1g; 3				

KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 RT Immunoglobulin domain; Alternative splicing; Polymorphism;
 FT SIGNAL 1 21
 FT CHAIN 22 374
 FT DOMAIN 22 292
 FT TRANSMEM 23 313
 FT DOMAIN 314 374
 FT DOMAIN 36 92
 FT DOMAIN 117 175
 FT DOMAIN 205 267
 FT DISULFID 43 85
 FT DISULFID 124 168
 FT DISULFID 212 260
 FT CARBOHYD 59 78
 FT CARBOHYD 78 152
 FT CARBOHYD 152 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT VARSPLIT 333 374
 FT VARIANT 26 26
 FT VARIANT 59 59
 FT CONFLICT 25 25
 FT SEQUENCE 374 AA; 42605 MW; 202AA8103ECP1686 CRC64;

Query Match Best Local Similarity 89.1%; Score 57; DB 1; Length 374;
 Matches 11; Conservative 91.7%; Pred. No. 0.0017;
 Mismatches 0; Indels 1; Gaps 0;

OY 1 EDSGLYCEAAT 12
 DB 253 EDSGLYCEAAT 264

RESULT 2
 CONT_CHICK STANDARD; PRT: 1010 AA.
 ID CONT_CHICK STANDARD; PRT: 1010 AA.
 AC P14781; P10450; (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CONTACTIN PRECURSOR (NEURAL CELL RECOGNITION MOLECULE F11).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90180453; PubMed=2627374;
 RA Bruemendorf T., Wolff J.M., Rainer F., Rathjen F.G.;
 RT "Neural cell recognition molecule F11: homology with fibronectin type
 RL III and immunoglobulin type C domains."; Neuron 2:1351-1361(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=WHITE LEHORN;
 RA Ranscht B., Dours M.T.;
 RT Sequence of contactin, a 130-kD glycoprotein concentrated in areas
 of interneuronal contact, defines a new member of the immunoglobulin
 RT supergene family in the nervous system."; J. Cell Biol. 107:1561-1573(1988).
 RN [3]
 RP GPI-ANCHOR
 RX MEDLINE=89286606; PubMed=2735929;
 RA Wolff J.M., Bruemendorf T., Rathjen F.G.;

RT "Neural cell recognition molecule F11: membrane interaction by
 covalently attached phosphatidylinositol."; Biochem. Biophys. Res. Commun. 161:931-938(1989).
 RL FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS
 CC SYSTEM DEVELOPMENT.
 CC -1 SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
 CC -1 SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1 SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN IN THE C-TERMINUS
 CC AND IS LONGER DUE TO A FRAMESHIFT.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X14877; CA33018.1; -
 DR EMBL: Y00813; CA68753.1; ALT_FRAME.
 DR PIR: J00094; J00094.
 DR PIR: S01998; S01998.
 DR HSSP: P20241; ICFB.
 DR InterPro: IPR001777;
 DR InterPro: IPR003006;
 DR Pfam: PF00041; fn3; 4.
 DR Pfam: PF00047; Ig; 6.
 DR Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
 KW Cell adhesion; Repeat;

FT CHAIN 1 19
 FT PROPEP 20 1010
 FT DOMAIN 50 113
 FT DOMAIN 143 210
 FT DOMAIN 247 308
 FT DOMAIN 336 389
 FT DOMAIN 420 482
 FT DOMAIN 510 581
 FT DOMAIN 593 599
 FT DOMAIN 600 701
 FT DOMAIN 702 803
 FT DOMAIN 804 900
 FT DOMAIN 901 996
 FT CARBOHYD 200 200
 FT CARBOHYD 249 249
 FT CARBOHYD 329 329
 FT CARBOHYD 448 448
 FT CARBOHYD 464 464
 FT CARBOHYD 485 485
 FT CARBOHYD 512 512
 FT CARBOHYD 582 582
 FT CARBOHYD 924 924
 FT SEQUENCE 1010 AA; 112507 MW; 2E38F071AE423AE1 CRC64;

Query Match Best Local Similarity 75.0%; Score 48; DB 1; Length 1010;
 Matches 8; Conservative 80.0%; Pred. No. 0.23;
 Mismatches 1; Indels 1; Gaps 0;

OY 1 EDSGLYCEA 10
 DB 294 EDSGLYCEA 303

RESULT 3
 CONT_MOUSE STANDARD; PRT: 1020 AA.
 ID CONT_MOUSE STANDARD; PRT: 1020 AA.
 AC P12960;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CONTACTIN PRECURSOR (NEURAL CELL SURFACE PROTEIN F3).

GN CNTNL.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=89340657; PubMed=2474555;
 RA Genarini G., Gibelli G., Rougon G., Mattei M.-G., Goidis C.;
 RT "The mouse neuronal cell surface protein F3: a phosphatidylinositol-
 anchored member of the immunoglobulin superfamily related to chicken
 contactin";
 RT contactin";
 RL J. Cell Biol. 109:775-788(1989).
 CC -1- FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS
 CC SYSTEM DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
 CC -1- MISCELLANEOUS: F3 SHARES WITH L1, N-CAM, MAG, AND OTHER CELL
 CC ADHESION MOLECULES FROM NERVOUS TISSUE THE L2/HNK-1 CARBOHYDRATE
 CC EPITOPE.
 CC -1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 DR EMBL: X14943; CAA33075.1; -
 DR PIR: S05944; S05944.
 DR HSSP: P40189; IBOU.
 DR MGD: MG1:105980; Cntnl.
 DR InterPro: IPR001777; -
 DR InterPro: IPR003006; -
 DR Pfam: PF00041; fn3; 4.
 DR Pfam: PF00047; Ig; 6.
 KW Immunoglobulin domain; glycoprotein; signal; GPI-anchor;
 KW Cell adhesion; Repeat.
 FT SIGNAL 1 20
 FT CHAIN 1 20
 FT PROPEP 21 1020
 FT DOMAIN 58 121
 FT DOMAIN 151 218
 FT DOMAIN 256 317
 FT DOMAIN 345 398
 FT DOMAIN 429 491
 FT DOMAIN 519 592
 FT DOMAIN 604 611
 FT DOMAIN 611 712
 FT DOMAIN 713 814
 FT DOMAIN 815 910
 FT DOMAIN 911 1006
 FT DISULFID 65 114
 FT DISULFID 158 211
 FT DISULFID 263 310
 FT DISULFID 352 391
 FT DISULFID 436 484
 FT DISULFID 526 585
 FT CARBOHYD 208 208
 FT CARBOHYD 258 258
 FT CARBOHYD 338 338
 FT CARBOHYD 457 457
 FT CARBOHYD 473 473
 FT CARBOHYD 494 494
 FT CARBOHYD 521 521
 FT CARBOHYD 593 593
 FT CARBOHYD 935 935
 SQ SEQUENCE 1020 AA; 11338 MW; 9DCDA40BAACBC7 CRC64;

Query Match 75.0%; Score 48; DB 1; Length 1020;
 Best Local Similarity 80.0%; Pred. No. 0.23;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EDSGLYCEA 10
 11111111
 Db 303 EDEGLYCEA 312
 RESULT 4
 DCC_HUMAN
 ID DCC_HUMAN STANDARD; PRT; 1447 AA.
 AC PA3146;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TUMOR SUPPRESSOR PROTEIN DCC PRECURSOR (COLORECTAL CANCER SUPPRESSOR).
 GN DCC
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95011532; PubMed=7926722;
 RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,
 RA Vogelstein B.;
 RT "The DCC gene product in cellular differentiation and colorectal
 RT tumorigenesis";
 RL Genes Dev. 8:1174-1183(1994).
 RN [2]
 RP SEQUENCE OF 1-750 FROM N.A.
 RX MEDLINE=90100559; PubMed=2294591;
 RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,
 RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
 RA Vogelstein B.;
 RT "Identification of a chromosome 18q gene that is altered in
 RT colorectal cancers";
 RL Science 247:49-56(1990).
 RN [3]
 RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS).
 RX MEDLINE=91121517; PubMed=1991322;
 RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,
 RA Oliner J.D., Kinzler K.W., Vogelstein B.;
 RT "Scrambled exons";
 RL Cell 64:607-613(1991).
 RN [4]
 RP GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.
 RX MEDLINE=94245241; PubMed=8188295;
 RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,
 RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;
 RT "The DCC gene: structural analysis and mutations in colorectal
 RT carcinomas";
 RL Genomics 19:525-531(1994).
 RN [5]
 RP VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.
 RX MEDLINE=94243823; PubMed=8187090;
 RA Miyake S., Nagai K., Yoshino K., Oto M., Endo M., Yuasa Y.;
 RT "Point mutations and allelic deletion of tumor suppressor gene DCC in
 RT human esophageal squamous cell carcinomas and their relation to
 RT metastasis";
 RL Cancer Res. 54:3007-3010(1994).
 RL
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.
 CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC
 CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS
 CC METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X76132; CAAS5735.1; -
 DR EMBL: M32292; AAA5751.1; -
 DR EMBL: M32286; AAA52174.1; -
 DR EMBL: M32288; AAA52175.1; ALT_SEQ.
 DR EMBL: M32290; AAA52176.1; -
 DR EMBL: M36696; AAA52177.1; -
 DR EMBL: M63700; AAA52178.1; -
 DR EMBL: M63702; AAA52179.1; -
 DR EMBL: M63718; AAA52180.1; -
 DR EMBL: M63698; AAA52181.1; -
 DR PIR: A54100; A54100.
 DR PIR: A40098; A40098.
 DR PIR: A38442; A38442.
 DR HSSP: P56276; ITLK.
 DR MIM: 120470; -
 DR InterPro: IPR001777; -
 DR InterPro: IPR003006; -
 DR Pfam: PF00041; fn3; 6.
 DR Pfam: PF00047; fn3; 4.
 DR PRINTS: PR00014; FNTYPEIII.
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
 KW Anti-oncogene; Disease mutation; Polymorphism.
 FT SIGNAL 1 25
 FT CHAIN 26 1447
 FT TRANSMEM 26 1097
 FT DOMAIN 1098 1122
 FT DOMAIN 1123 1447
 FT DOMAIN 54 124
 FT DOMAIN 154 219
 FT DOMAIN 254 317
 FT DOMAIN 345 407
 FT DOMAIN 426 522
 FT DOMAIN 525 618
 FT DOMAIN 619 716
 FT DOMAIN 722 816
 FT DOMAIN 840 940
 FT DOMAIN 941 1042
 FT DISULFID 61 117
 FT DISULFID 161 212
 FT DISULFID 261 310
 FT DISULFID 352 400
 FT CARBOHYD 299 399
 FT CARBOHYD 318 318
 FT CARBOHYD 478 478
 FT CARBOHYD 628 628
 FT CARBOHYD 702 702
 FT VARIANT 168 168
 FT VARIANT 201 201
 FT VARIANT 1375 1375
 FT CONFLICT 138 138
 FT CONFLICT 233 329
 FT CONFLICT 421 421
 FT SEQUENCE 1447 AA; 158456 MW; 4A861276ED0471F CRC64;

Query Match 73.4%; Score 47; DB 1; Length 1447;
 Best Local Similarity 80.0%; Pred. No. 0.51;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 111 DEGLYCEAS 120

RESULT 5
 DCC_MOUSE
 ID DCC_MOUSE STANDARD; PRT; 1447 AA.
 AC P70211;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TUMOR SUPPRESSOR PROTEIN DCC PRECURSOR.
 GN DCC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SOURCE FROM N.A.
 RC STRAIN=BALE/C; TISSUE=Brain;
 RX MEDLINE=96112625; Pubmed=8570174;
 RA Cooper H.M., Ames P., Britto J., Gad J., Wilks A.F.;
 RT Cloning of the mouse homologue of the deleted in colorectal cancer
 RT gene (MDC) and its expression in the developing mouse embryo.";
 RL Oncogene 11:2243-2254(1995).
 RN [2]
 RP REVISIONS.
 RC STRAIN=BALE/C; TISSUE=Brain;
 RA Cooper H.M.;
 RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THE PROTEIN ARE PRODUCED FROM
 CC THE SAME GENE BY THE USE OF ALTERNATIVE INITIATION SITES. A THIRD
 CC FORM WHICH IS EXPRESSED ONLY IN THE EMBRYO IS PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
 CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
 CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
 CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS
 CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION
 CC AND REMAIN AT THIS LEVEL IN THE ADULT.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X85788; CAAS9786.1; -
 DR HSSP: P56276; ITLK.
 DR MCD: MGI:94869; DCC.
 DR InterPro: IPR001777; -
 DR InterPro: IPR003006; -
 DR Pfam: PF00041; fn3; 6.
 DR PRINTS: PR00014; FNTYPEIII.
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
 KW Anti-oncogene; Alternative initiation; Alternative splicing.
 FT SIGNAL 1 25
 FT CHAIN 26 1447
 FT CHAIN 85 1447
 FT INIT_MET 85 85
 FT DOMAIN 26 1097
 FT TRANSMEM 1098 1122
 FT DOMAIN 1123 1447
 FT DOMAIN 54 124

ISOFORM.
 TUMOR SUPPRESSOR PROTEIN DCC, LONG
 ISOFORM.
 TUMOR SUPPRESSOR PROTEIN DCC, SHORT
 ISOFORM.
 FOR SHORT ISOFORM.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 IG-LIKE C2-TYPE DOMAIN.

FT DOMAIN 154 219 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 254 317 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 345 407 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 426 522 FIBRONECTIN TYPE-III.
 FT DOMAIN 525 618 FIBRONECTIN TYPE-III.
 FT DOMAIN 619 716 FIBRONECTIN TYPE-III.
 FT DOMAIN 722 816 FIBRONECTIN TYPE-III.
 FT DOMAIN 840 940 FIBRONECTIN TYPE-III.
 FT DOMAIN 941 1042 FIBRONECTIN TYPE-III.
 FT DISULFID 61 117 BY SIMILARITY.
 FT DISULFID 161 212 BY SIMILARITY.
 FT DISULFID 261 310 BY SIMILARITY.
 FT DISULFID 352 400 BY SIMILARITY.
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 702 702 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 819 838 MISSING (IN EMBRYONIC ISOFORM).
 SQ SEQUENCE 1447 AA; 158298 MW; 0D1F1097C22D5B9F CRC64;

Query Match 73.4%; Score 47; DB 1; Length 1447;
 Best Local Similarity 80.0%; Pred. No. 0.51;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 DSGLYOCEAA 11
 DB 111 DSGLYOCEAS 120

RESULT 6
 FCGL_MOUSE
 ID FCGL_MOUSE STANDARD; PRT; 404 AA.
 AC P26151;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I).
 GN FCGR1 OR FCGL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90111035; PubMed=2136886;
 RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RT "Molecular cloning and expression of the mouse high affinity Fc receptor for IgG.";
 RT J. Immunol. 144:371-378(1990).
 RL [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=9216339; PubMed=1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RT "Structure and mapping of the gene encoding mouse high affinity Fc gamma RI and chromosomal location of the human Fc gamma RI gene.";
 RT J. Immunol. 148:1570-1575(1992).
 RL [2]
 RN [1]
 RP FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN GAMMA. HIGH AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M31314; AAA40056.1;
 DR PIR: A43511; A43511.
 DR PIR: A46480; A46480.
 DR HSSP: P12319; 1ALT.
 DR MGD: MGI:95498; Fcgr1.
 DR InterPro: IPR03006;
 DR Pfam: PF00047; Ig_3.
 KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
 FT SIGNAL 1 24
 FT CHAIN 25 404
 FT FT
 FT FT
 FT DOMAIN 25 297 HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I.
 FT TRANSMEM 298 320 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 321 404 POTENTIAL.
 FT DOMAIN 46 66 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 67 124 IG-LIKE C2-TYPE DOMAIN (TRUNCATED).
 FT DOMAIN 154 216 IG-LIKE C2-TYPE DOMAIN.
 FT CARBOHYD 28 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 404 AA; 44887 MW; 1C4F0033842767E7 CRC64;

Query Match 71.9%; Score 46; DB 1; Length 404;
 Best Local Similarity 66.7%; Pred. No. 0.22;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDSGLYOCEAT 12
 DB 262 EDSGLYOCEVAT 273

RESULT 7
 CONT_HUMAN
 ID CONT_HUMAN STANDARD; PRT; 1018 AA.
 AC 012860; 012861; 014030;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CONTRACTIN PRECURSOR (GLYCOPROTEIN GP135).
 GN CNTN1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=95046335; PubMed=7959734;
 RA Berglund E.O., Ranscht B.;
 RT "Molecular cloning and in situ localization of the human contractin gene (CNTN1) on chromosome 12q11-q12.";
 RT Genomics 21:571-582(1994).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=94217459; PubMed=8164510;
 RA Reid R.A., Hemperly J.J.;
 RT "Identification and characterization of the human cell adhesion molecule contractin.";
 RT Brain Res. Mol. Brain Res. 21:1-8(1994).
 RL [2]
 RN [1]
 RP FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS SYSTEM DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U07819; AA67920.1; -
 DR EMBL: U07820; AA67921.1; -
 DR EMBL: 221488; CA679696.1; -
 DR HSP: P40189; IBOU.
 DR MIM: 600016; -
 DR InterPro: IPR001777; -
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; fn3; 4.
 DR Pfam: PF00047; 19; 6.
 KW Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
 KW Cell adhesion; Repeat; Alternative splicing.

FT CHAIN 1 20
 FT PROPEP 21 1018
 FT DOMAIN 58 121
 FT DOMAIN 151 218
 FT DOMAIN 256 317
 FT DOMAIN 345 398
 FT DOMAIN 429 491
 FT DOMAIN 519 590
 FT DOMAIN 602 609
 FT DOMAIN 711 812
 FT DOMAIN 813 908
 FT DOMAIN 909 1004
 FT DISULFID 158 211
 FT DISULFID 263 310
 FT DISULFID 352 391
 FT DISULFID 436 484
 FT DISULFID 526 583
 FT CARBOHYD 208 208
 FT CARBOHYD 258 258
 FT CARBOHYD 338 338
 FT CARBOHYD 457 457
 FT CARBOHYD 473 473
 FT CARBOHYD 494 494
 FT CARBOHYD 521 521
 FT CARBOHYD 591 591
 FT CARBOHYD 933 933
 FT VARSPIC 21 31
 FT CONFLICT 798 798
 SQ SEQUENCE 1018 AA; 113320 MW; 4B8FDC5BBD434ED5 CRC64;

Query Match 71.9%; Score 46; DB 1; Length 1018;
 Best Local Similarity 70.0%; Pred. No. 0.55;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDSGLYOCFA 10
 DB 303 EDEGIVECEA 312

RESULT 8
 AXOL-CHICK STANDARD; PRT: 1036 AA.
 AC P28685;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE AXONIN-1-PRECURSOR
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Archosauroida; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_TaxID=9031;
 CC [1]
 CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC TISSUE-Brain;
 CC MEDLINE=92174898; PubMed=1311675;
 CC RA Zueligg R.A., Rader C., Schroeder A., Kalousek M.B.,
 CC von Bohlen und Halbach F., Osterwalder T., Ivan C., Stoeckli E.T.,
 CC Affolter H., U., Fritz A., Hafen E., Sonderegger P.;
 CC "The axonally secreted cell adhesion molecule, axonin-1. Primary
 CC structure, immunoglobulin-like and fibronectin-type-III-like domains
 CC and glycosyl-phosphatidylinositol anchorage.";
 CC Eur. J. Biochem. 204:453-463(1992).
 CC -1- FUNCTION: AXON-ASSOCIATED CELL ADHESION MOLECULE (AXCAM) WHICH
 CC PROMOTES NEURITE OUTGROWTH BY INTERACTION WITH THE AXCAM L1 (G4)
 CC OF NEURITIC MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE NEURONAL MEMBRANE BY A
 CC GPI-ANCHOR.
 CC -1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X63101; CAA44815.1; -
 DR PIR: S22128; S22128.
 DR PIR: S22383; S22383.
 DR HSP: P56276; ITLK.
 DR InterPro: IPR001777; -
 DR InterPro: IPR003006; -
 DR Pfam: PF00041; fn3; 3.
 DR Pfam: PF00047; 19; 6.
 KW Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
 KW Cell adhesion; Repeat.

FT CHAIN 1 23
 FT PROPEP 24 1036
 FT DOMAIN 49 113
 FT DOMAIN 143 211
 FT DOMAIN 249 308
 FT DOMAIN 336 397
 FT DOMAIN 428 490
 FT DOMAIN 518 589
 FT DOMAIN 599 608
 FT DOMAIN 601 607
 FT DOMAIN 709 709
 FT DOMAIN 710 811
 FT DOMAIN 812 912
 FT DOMAIN 913 1009
 FT MOD. RES 224 224
 FT CARBOHYD 71 71
 FT CARBOHYD 199 199
 FT CARBOHYD 456 456
 FT CARBOHYD 472 472
 FT CARBOHYD 493 493
 FT CARBOHYD 520 520
 FT CARBOHYD 770 770
 FT CARBOHYD 900 900
 FT CARBOHYD 914 914
 SQ SEQUENCE 1036 AA; 113301 MW; 08B80143B879794 CRC64;

Query Match 71.9%; Score 46; DB 1; Length 1036;
 Best Local Similarity 80.0%; Pred. No. 0.56;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDSGLYOCFA 10

Cell adhesion; Repeat.
 KW SIGNAL 1 30
 FT CHAIN 31 21015
 FT PROPER 21016 1040
 FT DOMAIN 56 120
 FT DOMAIN 150 218
 FT DOMAIN 256 315
 FT DOMAIN 343 404
 FT DOMAIN 435 497
 FT DOMAIN 525 596
 FT DOMAIN 608 614
 FT DOMAIN 613 708
 FT DOMAIN 716 811
 FT DOMAIN 818 910
 FT DOMAIN 911 1005
 FT SITE 796 798
 FT CARBOHYD 78 78
 FT CARBOHYD 200 200
 FT CARBOHYD 206 206
 FT CARBOHYD 463 463
 FT CARBOHYD 479 479
 FT CARBOHYD 500 500
 FT CARBOHYD 527 527
 FT CARBOHYD 777 777
 FT CARBOHYD 832 832
 FT CARBOHYD 920 920
 FT CARBOHYD 942 942
 SQ SEQUENCE 1040 AA; 113042 MW; 6E707EF6614CB4B CRC64;

Query Match 71.98; Score 46; DB 1; Length 1040;
 Best Local Similarity 80.08; Pred. No. 0.57;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EDSGLYCCEA 10
 |||||
 Db 390 EDSGMTCVA 399

RESULT 11
 VCAM_MOUSE STANDARD; PRT: 739 AA.
 ID VCAM_MOUSE
 AC P29533;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE VASCULAR CELL ADHESION PROTEIN 1 PRECURSOR (V-CAM 1).
 GN VCAM1 OR VCAM-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB; TISSUE=Lung;
 RX MEDLINE=92181437; PubMed=1371918;
 RA Hession C., Moy P., Tizard R., Chisholm P., Williams C., Wysk M.,
 Burdick L., Miyake K., Kincade P., Lobb R.;
 RT "Cloning of murine and rat vascular cell adhesion molecule-1";
 RL Biochem. Biophys. Res. Commun. 183:163-169(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph node;
 RX MEDLINE=93246254; PubMed=7683304;
 RA Araki M., Araki K., Vassalli P.;
 RT "Cloning and sequencing of mouse VCAM-1 cDNA";
 RL Gene 126:261-264(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=Embryo;
 RX MEDLINE=94117008; PubMed=7507076;
 RA Cypulsky M.I., Allan-Motamed M., Collins T.;
 RT "Structure of the murine VCAM1 gene";

RL Genomics 18:387-391(1993).
 RN [4]
 RP SEQUENCE OF 1-693 FROM N.A.
 RC STRAIN=NIH SWISS, AND 129/SV;
 RA Kumar A.G., Dai Y.X., Kozak C.A., Mims M.P., Gatto A.M. Jr.,
 Ballantyne C.M.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 1-345 FROM N.A. (GPI-ANCHORED ISOFORM).
 RC STRAIN=FVB; TISSUE=Lung;
 RX MEDLINE=93232042; PubMed=7682556;
 RA Moy P., Lobb R., Tizard R., Olson D., Hession C.;
 RT "Cloning of an inflammation-specific phosphatidylinositol-linked
 form of murine vascular cell adhesion molecule-1";
 RL J. Biol. Chem. 268:8835-8841(1993).
 RN [6]
 RP SEQUENCE OF 1-345 FROM N.A. (GPI-ANCHORED ISOFORM).
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RX MEDLINE=95015899; PubMed=7523515;
 RA Kumar A.G., Dai Y.X., Kozak C.A., Mims M.P., Gatto A.M.,
 Ballantyne C.M.;
 RT "Murine VCAM-1. Molecular cloning, mapping, and analysis of a
 truncated form";
 RL J. Immunol. 153:4088-4098(1994).
 RN [7]
 RP SEQUENCE OF 311-345 FROM N.A. (GPI-ANCHORED ISOFORM).
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RX MEDLINE=93317595; PubMed=7687058;
 RA Terry R.W., Kwee L., Levine J.F., Labow M.A.;
 RT "Cytokine induction of an alternatively spliced murine vascular cell
 adhesion molecule (VCAM) mRNA encoding a
 glycosylphosphatidylinositol anchored VCAM protein";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:5919-5923(1993).
 RN [8]
 RP SEQUENCE OF 1-21 FROM N.A.
 RC TISSUE=Endothelial cells;
 RA Korenaga R., Ando J., Tsuboi H., Kamlya A.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
 IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
 INTEGRIN VLA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
 TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A
 PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
 EMIGRATION TO SITES OF INFLAMMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG ISOFORM) OR
 ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (SHORT ISOFORM).
 CC -1- ALTERNATIVE PRODUCTS: THE GPI-ANCHORED AND THE TRANSMEMBRANE
 ISOFORMS ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON INFLAMED VASCULAR ENDOTHELIUM, AS
 WELL AS ON MACROPHAGE-LIKE AND DENDRITIC CELL TYPES IN BOTH NORMAL
 AND INFLAMED TISSUE.
 CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M84487; AAA40545.1;
 DR EMBL: X67783; CAA47989.1;
 DR EMBL: L22355; AAA16921.1;
 DR EMBL: L22301; AAA16921.1; JOINED.
 DR EMBL: L22349; AAA16921.1; JOINED.
 DR EMBL: L22350; AAA16921.1; JOINED.
 DR EMBL: L22351; AAA16921.1; JOINED.
 DR EMBL: L22352; AAA16921.1; JOINED.
 DR EMBL: L22353; AAA16921.1; JOINED.
 DR EMBL: L22354; AAA16921.1; JOINED.
 DR EMBL: L22350; AAA16920.1;
 DR EMBL: L22301; AAA16920.1; JOINED.

DR EMBL: L22349; AAA16920.1; JOINED.
 DR EMBL: U12878; AAB60659.1; ALT_SEQ.
 DR EMBL: U12879; AAB60660.1; ALT_SEQ.
 DR EMBL: U12880; AAB60661.1; ALT_SEQ.
 DR EMBL: U12874; AAB60662.1; ALT_SEQ.
 DR EMBL: U12871; AAB60663.1; ALT_SEQ.
 DR EMBL: U12883; AAB60664.1; ALT_SEQ.
 DR EMBL: U12881; AAB60665.1; ALT_SEQ.
 DR EMBL: U12882; AAB60666.1; ALT_SEQ.
 DR EMBL: U12875; AAB60667.1; ALT_SEQ.
 DR EMBL: U12872; AAB60668.1; ALT_SEQ.
 DR EMBL: U12876; AAB60669.1; ALT_SEQ.
 DR EMBL: U12873; AAB60670.1; ALT_SEQ.
 DR EMBL: U12877; AAB60671.1; ALT_SEQ.
 DR EMBL: L08431; AAA40546.1; -
 DR EMBL: U12884; AAA40547.1; -
 DR EMBL: U12541; AAC37607.1; -
 DR EMBL: U43327; AAB88576.1; -
 DR PIR: JS0674; JS0674.
 DR PIR: JN0581; JN0581.
 DR HSP: P19320; 1YCA.
 DR MGD; MGI:98926; Vcam1.
 DR InterPro: IPR003006; -
 KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
 KW GPI-anchor; Signal; Alternative splicing.
 FT SIGNAL 1 24
 FT CHAIN 25 739
 FT PROPEP 311 345
 FT LIPID 310 310
 FT TRANSSEM 699 720
 FT DOMAIN 721 739
 FT DOMAIN 129 38
 FT DOMAIN 129 91
 FT DOMAIN 129 198
 FT DOMAIN 129 287
 FT DOMAIN 129 379
 FT DOMAIN 129 418
 FT DOMAIN 129 496
 FT DOMAIN 129 525
 FT DOMAIN 129 575
 FT CARBOHYD 225 225
 FT CARBOHYD 273 273
 FT CARBOHYD 424 424
 FT CARBOHYD 531 531
 FT CARBOHYD 561 561
 FT VARSPLIC 310 345
 FT VARSPLIC 346 739
 FT CONFLICT 693 693
 FT SEQUENCE 739 AA; 81317 MM; 3D2134C341E5E449 CRC64;
 Query Match 70.3%; Score 45; DB 1; Length 739;
 Best Local Similarity 50.0%; Pred. No. 0.62;
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

DE VASCULAR CELL ADHESION PROTEIN 1 PRECURSOR (V-CAM 1).
 GN VCAM1 OR VCAM-1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Lung;
 RX MEDLINE=92181437; PubMed=1371918;
 RA Hession C., Moy P., Tizard R., Chisholm P., Williams C., Wyse M.,
 RA Burdick L., Miyake K., Kincaid P., Lobb R.;
 RT "Cloning of murine and rat vascular cell adhesion molecule-1";
 RL Biochem. Biophys. Res. Commun. 183:163-169(1992).
 CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
 CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
 CC INTEGRIN VIA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
 CC TRANSDUCTION. THE VCAM1/VIA4 INTERACTION MAY PLAY A
 CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
 CC EMIGRATION TO SITES OF INFLAMMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON INFLAMED VASCULAR ENDOTHELIUM, AS
 CC WELL AS ON MACROPHAGE-LIKE AND DENDRITIC CELL TYPES IN BOTH NORMAL
 CC AND INFLAMED TISSUE.
 CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M84488; AAA42332.1; -
 CC PIR: JS0675; JS0675.
 CC HSP: P19320; 1YCA.
 CC InterPro: IPR003006; -
 DR PIR: JS0675; JS0675.
 DR HSP: P19320; 1YCA.
 DR InterPro: IPR003006; -
 KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
 KW Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 739
 FT PROPEP 311 345
 FT LIPID 310 310
 FT TRANSSEM 699 720
 FT DOMAIN 721 739
 FT DOMAIN 129 38
 FT DOMAIN 129 91
 FT DOMAIN 129 198
 FT DOMAIN 129 287
 FT DOMAIN 129 379
 FT DOMAIN 129 418
 FT DOMAIN 129 496
 FT DOMAIN 129 525
 FT DOMAIN 129 575
 FT CARBOHYD 225 225
 FT CARBOHYD 273 273
 FT CARBOHYD 424 424
 FT CARBOHYD 531 531
 FT CARBOHYD 561 561
 FT VARSPLIC 310 345
 FT VARSPLIC 346 739
 FT CONFLICT 693 693
 FT SEQUENCE 739 AA; 81246 MM; 5C608E5A1A1B100C CRC64;
 Query Match 68.8%; Score 44; DB 1; Length 739;
 Best Local Similarity 50.0%; Pred. No. 0.96;
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

FT FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT VARSPLIC 694 837 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT
FT FT
FT FT
SO SEQUENCE 837 AA; 93203 MW; 70473B053A2D65A5 CRC64;
Query Match 68.8%; Score 44; DB 1; Length 837;
Best Local Similarity 72.7%; Pred. No. 1,1;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0
QY 2 DSGLYOCEAAT 12
    ||| |||||
Db 374 DSGRYDCEAAS 384

RESULT 14
NCM2_HUMAN
ID NCM2_HUMAN STANDARD; PRT; 837 AA.
AC 015394;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NEURAL CELL ADHESION MOLECULE 2 PRECURSOR (N-CAM 2).
GN NCAM2 OR NCAM21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCB1_TaxID=9606;
[1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97369930; PubMed=9226371;
RA Paoloni-Giacobino A., Chen H., Antonarakis S.E.;
RT "Cloning of a novel human neural cell adhesion molecule gene (NCAM2)
RT that maps to chromosome region 21q21 and is potentially involved in
RT down syndrome."
RL Genomics 43:43-51 (1997).
CC -!- FUNCTION: MAY PLAY IMPORTANT ROLES IN SELECTIVE FASCICULATION AND
CC -!- ZONE-TO-ZONE PROJECTION OF THE PRIMARY OLFACTORY AXONS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED MOST STRONGLY IN ADULT AND FETAL
CC BRAIN.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 FIBRINECTIN TYPE III-LIKE DOMAINS.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
-----
DR DR EMBL: U75330.1 ABB80803.1;
DR DR MIM: 602040;
DR DR InterPro: IPR001777;
DR DR InterPro: IPR003006;
DR DR Pfam: PF00041; fn3: 2;
DR DR Pfam: PF00047; 1g: 5;
DR DR Cell adhesion; Transmembrane; Glycoprotein; Repeat;
DR DR Immunoglobulin domain; Signal.
DR DR

```

Query Match	Best Local Similarity	67.2%;	Score 43;	DB 1;	Length 837;
Matches	6;	Conservative	4;	Mismatches	0;
				Indels	Gaps
Oy	1	EDSGLOCEA 10			
		:			
Db	86	EDAGYRCQA 95			
RESULT 15					
VG22 HUMAN	ID	VG22 HUMAN	STANDARD;	PRT;	1356 AA.
AC	P35968;	O60723;	Q14178;		
DT	01-JUN-1994	(Rel. 29, Created)			
DT	01-OCT-2000	(Rel. 40, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 2 PRECURSOR (EC 2.7.1.112)				
DE	(VEGFR-2) (KINASE INSERT DOMAIN RECEPTOR) (FLK-1).				
GN	KDR OR FLK1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
	[1]				
RA	SEQUENCE FROM N.A.				
RA	Yin L.Y., Mu Y., Patterson C.;				
RT	"Full length human KDR/Flk-1 sequence."				
RL	Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.				
	[2]				
RA	SEQUENCE FROM N.A.				
RA	TISSUE=Umbilical vein;				
RC	Yu Y., Whitney R.G., Sato J.D.;				
RT	"Coding region for human VEGF receptor KDR (VEGFR-2)."				
RL	Submitted (May-1998) to the EMBL/Genbank/DBJ databases.				
	[3]				
RA	SEQUENCE OF 3-1356 FROM N.A.				
RA	TISSUE=Umbilical vein;				
RX	MEDLINE=92019839; PubMed=1656371;				
RA	Terman B.L., Carrion M.E., Kovacs E., Rasmussen B.A., Eddy R.L.,				
RT	Shows T.B.;				
RT	"Identification of a new endothelial cell growth factor receptor				
RL	tyrosine kinase."				
RL	Oncogene 6:1677-1683(1991).				
	[4]				
RA	SEQUENCE OF 1-22 FROM N.A.				
RA	MEDLINE=96032749; PubMed=7559454;				
RX					

FT	CARBOHYD	374	374	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	395	395	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	511	511	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	523	523	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	580	580	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	613	613	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	619	619	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	631	631	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	675	675	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	704	704	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	721	721	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	MOD_RES	1059	1059	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).	
FT	CONFLICT	2	2	Q -> E (IN REF. 2).	
FT	CONFLICT	772	772	A -> T (IN REF. 3).	
FT	CONFLICT	787	787	R -> G (IN REF. 3).	
FT	CONFLICT	835	835	K -> N (IN REF. 3).	
FT	CONFLICT	848	848	V -> E (IN REF. 3).	
FT	CONFLICT	1347	1347	S -> T (IN REF. 3).	
SO	SEQUENCE	1356 AA;	151526 MW;	59E7C4B05CFEBB3	CRC64;

Query Match

67.28;

Score 43;

DB 1;

Length 1356;

Best Local Similarity 58.38;

Pred. No. 2.7;

Mismatches 3;

Indels 0;

Gaps 0;

QY 1 EDSGLYOCFAAT 12

Db 730 EDEGLYTCQACS 741

Search completed: June 4, 2001, 12:24:07
job time: 569 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:23:27 ; Search time 107.68 seconds
(without alignments)
13.062 Million cell updates/sec

Title: US-09-284-107-26
Perfect score: 64
Sequence: 1 EDGLGYCEAAT 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_15:
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	57	89.1	374	4	092663	092663 homo sapien
2	57	89.1	375	4	092495	092495 homo sapien
3	49	76.6	1232	13	090284	090284 carassius a
4	48	75.0	105	13	09YHT0	09YHT0 ginglymsto
5	48	75.0	555	13	090874	090874 gallus gall
6	48	75.0	1021	11	063198	063198 rattus norv
7	48	75.0	1256	11	035158	035158 rattus norv
8	47	73.4	412	11	063611	063611 rattus norv
9	47	73.4	412	11	09R1E1	09R1E1 rattus norv
10	47	73.4	415	11	060977	060977 mus musculi
11	47	73.4	1445	11	063155	063155 caenorhabdi
12	47	73.4	1287	5	018100	018100 caenorhabdi
13	46	71.9	234	4	09NP99	09NP99 homo sapien
14	46	71.9	1018	6	028106	028106 bos taurus
15	46	71.9	1040	13	09W675	09W675 brachydanto
16	46	71.9	1250	11	088971	088971 mus musculi
17	46	71.9	2294	5	09N3X8	09N3X8 caenorhabdi
18	45	70.3	472	11	09WVH7	09WVH7 mus musculi
19	45	70.3	475	11	062056	062056 mus musculi

20	45	70.3	475	11	P70161	P70161 mus musculi
21	45	70.3	476	11	09R038	09R038 mus musculi
22	45	70.3	1197	13	090478	090478 brachydanto
23	44	68.8	162	13	091561	091561 xenopus lae
24	44	68.8	175	5	09VFD9	09VFD9 drosophila
25	44	68.8	243	5	09VY33	09VY33 drosophila
26	44	68.8	267	13	090529	090529 ginglymsto
27	44	68.8	330	11	P97269	P97269 cavia porce
28	44	68.8	337	11	P97268	P97268 cavia porce
29	44	68.8	349	6	09M2T0	09M2T0 bos taurus
30	44	68.8	406	5	09VNB1	09VNB1 drosophila
31	44	68.8	471	11	09R1F9	09R1F9 mus musculi
32	44	68.8	515	11	09JKE6	09JKE6 mus musculi
33	44	68.8	516	11	09J117	09J117 mus musculi
34	44	68.8	561	11	062929	062929 rattus norv
35	44	68.8	739	11	063669	063669 rattus norv
36	44	68.8	858	5	018466	018466 hirudo medi
37	44	68.8	1427	13	091562	091562 xenopus lae
38	44	68.8	7962	4	010465	010465 homo sapien
39	43	67.2	100	13	09W628	09W628 ginglymsto
40	43	67.2	104	5	09U543	09U543 aedes aegy
41	43	67.2	123	13	090843	090843 gallus gall
42	43	67.2	136	13	09YHL6	09YHL6 ginglymsto
43	43	67.2	194	5	022125	022125 caenorhabdi
44	43	67.2	282	4	09UFM8	09UFM8 homo sapien
45	43	67.2	393	11	P97547	P97547 rattus norv

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	374 AA
092663	092663			
AC	092663			
DT	01-FEB-1997 (TREMREL. 02, Created)			
DT	01-FEB-1997 (TREMREL. 02, Last sequence update)			
DT	01-OCT-2000 (TREMREL. 15, Last annotation update)			
DE	FC GAMMA RECEPTOR I.			
GN	AI.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BLOOD;			
RX	MEDLINE=3055454; PubMed=1430234;			
RA	Porges A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,			
RA	Kimberly R.P.;			
RT	"Novel Fc gamma receptor I family gene products in human mononuclear			
RT	cells.";			
RL	J. Clin. Invest. 90:2102-2109(1992).			
DR	EMBL: L03418; AAA36049.1;			
DR	HSSP: P12319; IALT.			
DR	INTERPRO: IPR003006;			
DR	PIR: PFO0047; Iq; 3.			
DR	PRODO: PD002534; -; 1.			
SO	SEQUENCE 374 AA; 42632 MW; D33D59398CEA699 CRC64;			
Query Match	89.1%; Score 57; DB 4; Length 374;			
Best Local Similarity	91.7%; Pred. No. 0.0032;			
Matches 11; Conservative	0; Mismatches 1; Indels 0; Gaps 0;			
OY	1 EDGLGYCEAAT 12			
DB	253 EDGLGYCEAAT 264			
RESULT	2			
092495	PRELIMINARY;	PRT;	375 AA.	
ID	092495			

```

AC 092495; 01-FEB-1997. (TREMblrel 02. Created)
DT 01-FEB-1997 (TREMblrel 02. Last sequence update)
DT 01-OCT-2000 (TREMblrel 15. Last annotation update)
DE FC GAMMA RECEPTOR TYPE I (FC GAMMA RIB=FC GAMMA RECEPTOR).
GN CD64 OR FC<GAMMA>RIB.
OS Homo sapiens (Human).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eumaryota; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Blood;
RA Benech P.D., Sastry K.N., Iyer R.R., Eichbaum Q.G., Ravesh D.,
RA Ezekovitz A.B.,
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-224 FROM N.A.
RX MEDLINE=93018827; PubMed=1402657;
RA Benech P.D., Sastry K., Iyer R.R., Eichbaum Q.G., Ravesh D.P.,
RA Ezekovitz R.A.,
RT "Definition of Interferon gamma-response elements in a novel human Fc
RT gamma receptor gene (Fc gamma RIb) and characterization of the gene
RT structure."
RL J. Exp. Med. 176:1115-1123(1992).
DR EMBL; M91555; AAAS8414.1; -
DR EMBL; M91550; AAAS8414.1; JOINED
DR EMBL; M91551; AAAS8414.1; JOINED
DR EMBL; M91552; AAAS8414.1; JOINED
DR EMBL; M91553; AAAS8414.1; JOINED
DR EMBL; M91554; AAAS8414.1; JOINED
DR EMBL; S45709; AAD13842.1; JOINED
DR EMBL; S45707; AAD13842.1; JOINED
DR EMBL; S45708; AAD13842.1; JOINED
DR EMBL; S45704; AAD13842.1; JOINED
DR EMBL; S45705; AAD13842.1; JOINED
DR HSSP; P12319; 1ALT.
DR INTERPRO; IPRO03006; -
DR PFAM; PF00047; 1q; 3.
DR PRODOM; PD002534; -; 1
SO SEQUENCE 375 AA; 42881 MW; A84d464c70dd0p91 CRC64;

```

```

QY      1  EDSCLYGCEAAT 12
        ||||| |||||
Db      254  EDSCLYGCEAAT 265

Query Match      99.1%;  Score 57;  DB 4;  Length 375;
Best Local Similarity 91.7%;  Pred. No. 0.0032;
Matches 11;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0.

RESULT      3
ID          090284      PRELIMINARY;  PRT: 1232 AA.
AC          090284;

DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE L1-LIKE CELL ADHESION MOLECULE ANTIGEN E587.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OC NCBI_TaxID=9557;
RN 11
RP SEQUENCE FROM N.A.
RA Giordano S., Laessing U., Lottspeich F., Stuermer C.A.O.;
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; U55211; AAA99159.1; -
DR HSSP; P20241; ICFB
DR INTERPRO; IPR001777; -
DR INTERPRO; IPR003066; -

```

DR PEAM: PE00041; fn3; 5.
DR PEAM: PE00041; 19; 6.
DR PRINTS: PR00014; ENTPPEIT.
SQ SEQUENCE 1232 AA; 136482 MW; 43086C06862C3B55 CRC64

Query Match	76.6%;	Score 49;	DB 13;	Length 1332;
Best Local Similarity	66.7%;	Pred. No. 0.41;		
Matches	8;	Conservative	2;	Mismatches 2;
				Indels 0;
				Gaps 0;
Oy	1	EDSGYCEAAT	12	
DB	553	EDSGYCEAAT	564	

RESULT	4			
09YHTO				
ID	09YHTO	PRELIMINARY;	PRT;	105 AA.
AC	Q9YHT0;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, last annotation update)			
DE	ANTIGEN RECEPTOR (FRAGMENT).			
GN	NAR.			
OS	Ginglymostoma cirratum (Nurse shark).			
OC	Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;			
OC	Elasmobranchii; Galeomorphii; Galeoidea; Oreotolobiformes;			
OC	Ginglymostomidae; Ginglymostoma.			
OX	NCBI_TaxID=7801;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99045686; PubMed=9826702;			
RA	Diaz M., Greenberg A.S., Flajnik M.F.;			
RT	"Somitic hypermutation of the new antigen receptor gene (NAR) in the			
RT	nurse shark does not generate the repertoire: possible role in			
RT	antigen-driven reactions in the absence of germinal centers.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:14343-14348(1998).			
DR	EMBL: AF095951; AAC84071.1; -.			
DR	INTERPRO: IPR003006; -.			
DR	PFAM: PF00047; 1g; 1.			
KW	Receptor.			
FT	NON_TER	1	1	
FT	NON_TER	105	105	
SO	SEQUENCE	105 AA; 11437 MW;	5F8EB7EA306A553F CRC64;	

Query Match	75.0%	Score 48;	DB 13;	Length 105;
Best Local Similarity	66.7%	Pred. NO. 0.049;		
Matches 8;	Conservative	2;	Mismatches 0;	Gaps 0;
QY	1 EDSGLYOCENAT 12			
	:::			
Db	73 EDSCTYOCENVS 84			
RESULT 5				
ID	Q90874	PRELIMINARY;	PRT;	555 AA.
AC	Q90874;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	IL-1 RECEPTOR (PRECURSOR.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
NCBI	TaxID=9031;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=92175529; PubMed=1531799;			
RX				
RA	Guida S., Heguy A., Meili M.;			
RT	"The chicken IL-1 receptor: differential evolution of the cytoplasmic			

RT and extracellular domains."
 RL Gene.111.239-243(1992).
 DR EMBL; M81846; AAA48924.1; -.
 DR HSSP; P14778; 11TB.
 DR INTERPRO; IPR000157; -.
 DR INTERPRO; IPR003006; -.
 DR PFAM; PF00047; 1g; 3.
 DR PFAM; PF01582; TIR; 1.
 KW Signal; Transmembrane.
 FT SIGNAL 1 19
 FT CHAIN 20 555
 SQ SEQUENCE 555 AA; 63995 MW; 4DB56DD7B1B817AA CRC64;

Query Match 75.0%; Score 48; DB 13; Length 555;
 Best Local Similarity 88.9%; Pred. NO. 0.28;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDSGLYOCE 9
 DB 91 EDSGLYCE 99

RESULT 6
 063198 PRELIMINARY; PRT: 1021 AA.

AC 063198
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE RAT NEURAL ADHESION MOLECULE F3, COMPLETE CDS.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WTSTAR; TISSUE=BRAIN;
 RX MEDLINE=95295987; PubMed=7777204;
 RA Hosoya H., Shimazaki K., Kobayashi S., Takahashi H., Shirasawa T.,
 RA Takenawa T., Watanabe K.;
 RT "Developmental expression of the neural adhesion molecule F3 in the
 RT rat brain."
 RL Neurosci. Lett. 186:83-86(1995).
 DR EMBL; D38492; BAA07504.1; -.
 DR HSSP; P40189; 18OU.
 DR INTERPRO; IPR001777; -.
 DR INTERPRO; IPR003006; -.
 DR INTERPRO; IPR003006; -.
 DR PFAM; PF00047; fn3; 4.
 DR PFAM; PF00047; 1g; 6.
 DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
 SQ SEQUENCE 1021 AA; 113494 MW; FCBDC13055EE5C68 CRC64;

Query Match 75.0%; Score 48; DB 11; Length 1021;
 Best Local Similarity 80.0%; Pred. NO. 0.53;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDSGLYOCEA 10
 DB 303 EDSGLYCEA 312

RESULT 7
 035158 PRELIMINARY; PRT: 1256 AA.

AC 035158
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CDO.
 GN CDO.

OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97362072; PubMed=9214393;
 RA Kang J.S., Gao M., Feinleib J.L., Cotter P.D., Guadagno S.N.,
 RA Krauss R.S.;
 RT "CDO: an oncogene-, serum-, and anchorage-regulated member of the
 RT Ig/fibronectin type III repeat family."
 RL J. Cell Biol. 138:203-213(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Krauss R.S., Kang J.S., Gao M., Feinleib J.L.,
 RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Krauss R.S., Kang J.S., Gao M., Feinleib J.L.,
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF004840; AAC34735.1; -.
 DR HSSP; P56276; 1TLK.
 DR INTERPRO; IPR001777; -.
 DR INTERPRO; IPR003006; -.
 DR PFAM; PF00047; fn3; 3.
 DR PFAM; PF00047; 1g; 5.
 SQ SEQUENCE 1256 AA; 136203 MW; 775805754F0C22EA CRC64;

Query Match 75.0%; Score 48; DB 11; Length 1256;
 Best Local Similarity 90.0%; Pred. NO. 0.65;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDSGLYOCEA 10
 DB 373 EDSGLYCEA 382

RESULT 8
 063611 PRELIMINARY; PRT: 412 AA.

AC 063611
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE TUMOR-ASSOCIATED ANTIGEN PRECURSOR.
 GN PE4.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BDIX; TISSUE=COLON TUMOR;
 RX MEDLINE=94253144; PubMed=8195207;
 RA Chadenau C., Lekomilac B., Denis M.G.,
 RA "A novel member of the immunoglobulin gene superfamily expressed in
 RT rat carcinoma cell lines."
 RL J. Biol. Chem. 269:15601-15605(1994).
 DR EMBL; L12025; AAB80767.1; -.
 DR INTERPRO; IPR003006; -.
 DR PFAM; PF00047; 1g; 3.
 KW Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 412
 SQ SEQUENCE 412 AA; 44509 MW; 98025D4C2161859B CRC64;

Query Match 73.4%; Score 47; DB 11; Length 412;
 Best Local Similarity 58.3%; Pred. NO. 0.32;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDSGLYOCEAAT 12

DB 123 EDEGIECOIAT 134

RESULT 9

Q9RIE1 PRELIMINARY; PRT; 412 AA.

AC Q9RIE1
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE TUMOR-ASSOCIATED GLYCOPROTEIN E4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BN/SSNHS;
RA Baurty B., Masson D., Lustenberger P., Denis M.G.;
RT "Structure of the rat Tagged gene."
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF125562; AAD25486.1; JOINED.
DR EMBL; AF125555; AAD25486.1; JOINED.
DR EMBL; AF125556; AAD25486.1; JOINED.
DR EMBL; AF125557; AAD25486.1; JOINED.
DR EMBL; AF125558; AAD25486.1; JOINED.
DR EMBL; AF125559; AAD25486.1; JOINED.
DR EMBL; AF125560; AAD25486.1; JOINED.
DR EMBL; AF125561; AAD25486.1; JOINED.
DR HSSP; Q13740; 1KJC.
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 1g; 3.
SQ SEQUENCE 412 AA; 44483 MW; D57650382569F093 CRC64;

Query Match 73.4%; Score 47; DB 11; Length 412;

Best Local Similarity 58.3%; Pred. No. 0.32;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDSGLYOCEAT 12
DB 123 EDEGIECOIAT 134

RESULT 10
Q60977 PRELIMINARY; PRT; 415 AA.
AC Q60977
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE TUMOR-ASSOCIATED GLYCOPROTEIN E4.
OS TAAL OR TAGER.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA MEDLINE-96372659; PubMed-8678999;
RA Chadenou C., Lecomille B., Lecabellier M., Mattei M., Meflah K.,
RA Denis M.G.;
RT Isolation and chromosomal location of me4, a novel murine gene of the
RT immunoglobulin superfamily *;
RL Mamm. Genome 7:636-637(1996).
DR EMBL; U35836; AAB17502.1;
DR HSSP; Q13740; 1KJC.
DR MGD; MGI:107741; Taal.
DR INTERPRO; IPR001064;
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 1g; 2.

DR PROSITE; PS00225; CRYSTALLIN BETAGAMMA; UNKNOWN.1.
SQ SEQUENCE 415 AA; 45280 MW; 159AF8D6DEF90DOF CRC64;

Query Match 73.4%; Score 47; DB 11; Length 415;

Best Local Similarity 58.3%; Pred. No. 0.32;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDSGLYOCEAT 12
DB 123 EDEGIECOIAT 134

RESULT 11

Q18100 PRELIMINARY; PRT; 1287 AA.

AC Q18100
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE C18F3.2 PROTEIN.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium *;
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Du Z., Gattung S.;
RT "The sequence of C. elegans cosmid C18F3 *;
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; U50067; AAA93439.2;
DR HSSP; P20241; 1CFB.
DR INTERPRO; IPR001777;
DR INTERPRO; IPR003006;
DR PFAM; PF00041; fn3; 6.
DR PFAM; PF00047; 1g; 6.
SQ SEQUENCE 1287 AA; 144194 MW; 935460E2707CE0B3 CRC64;

Query Match 73.4%; Score 47; DB 5; Length 1287;

Best Local Similarity 88.9%; Pred. No. 1.1;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DSGLYOCEA 10
DB 168 DSGLYOCEA 176

RESULT 12

Q63155 PRELIMINARY; PRT; 1445 AA.

AC Q63155
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE COLORECTAL TUMOR SUPPRESSOR.
 GN DCC.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97015074; PubMed=8861902;
 RA Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S.Y.,
 RA Culicetti J.G., Tessier-Lavigne M.;
 RT Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.*;
 RL Cell 87:175-185(1996).
 RN [2]
 RP SEQUENCE OF 387-420 FROM N.A.
 RX MEDLINE=90100559; PubMed=2294591;
 RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,
 RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
 RA Vogelstein B.;
 RT Identification of a chromosome 18q gene that is altered in colorectal
 RT cancers.*;
 RL Science 247:49-56(1990).
 DR EMBL: U68725; AAB41099.1; -;
 DR EMBL: M32291; AAA41086.1; -;
 DR HSSP: P56276; ITLK.
 DR INTERPRO: IPR001777; -;
 DR INTERPRO: IPR003006; -;
 DR PFAM: PF00041; fn3; 6.
 DR PFAM: PF00047; 19; 4.
 DR PRINTS: PRO0014; FMTYPE11.
 SQ SEQUENCE 1445 AA; 157940 MW; 084F625954481988 CRC64;

Query Match 73.4%; Score 47; DB 11; Length 1445;
 Best Local Similarity 80.0%; Pred. No. 1.2;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2_DSGLYCEAA 11
 Db 111 DEGLYCEAS 120

RESULT 13
 O9NP99 PRELIMINARY; PRT; 234 AA.
 AC O9NP99;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE TRIGGERING RECEPTOR EXPRESSED ON MONOCYTES 1.
 GN TREM1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20261626; PubMed=10799849;
 RA Bouchon A., Dietrich J., Colonna M.;
 RT Cutting edge: inflammatory responses can be triggered by TREM-1, a
 RT novel receptor expressed on neutrophils and monocytes.*;
 RL J. Immunol. 164:4991-4995(2000).
 DR EMBL: AF287008; AAF90197.1; -;
 DR EMBL: AF196329; AAF71694.1; -;
 KW Receptor.

SQ SEQUENCE 234 AA; 26386 MW; AAL14696E35D4D45 CRC64;

Query Match 71.9%; Score 46; DB 4; Length 234;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1_EDSGLYOC 8
 Db 106 EDSGLYOC 113

RESULT 14
 O28106 PRELIMINARY; PRT; 1018 AA.
 AC O28106;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DE F3/F11/CONTACTIN PRECURSOR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN.
 RX MEDLINE=95369697; PubMed=7642103;
 RA Matanabe K., Shimazaki K., Hosoya H., Fukumachi F., Takenawa T.;
 RT Cloning of the cDNA encoding neural adhesion molecule F3 from bovine
 RT brain.*;
 RL Gene 160:245-248(1995).
 DR EMBL: D32135; BAA06861.1; -;
 DR INTERPRO: IPR001777; -;
 DR INTERPRO: IPR002052; -;
 DR INTERPRO: IPR003006; -;
 DR PFAM: PF00041; fn3; 4.
 DR PFAM: PF00047; 19; 6.
 DR PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
 KW Signal; Cell adhesion.
 FT SIGNAL 1 20
 SQ SEQUENCE 1018 AA; 113384 MW; 551FC10ED7920341 CRC64;

Query Match 71.9%; Score 46; DB 6; Length 1018;
 Best Local Similarity 70.0%; Pred. No. 1.3;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1_EDSGLYCEA 10
 Db 303 EDEGLYCEA 312

RESULT 15
 O9W675 PRELIMINARY; PRT; 1040 AA.
 AC O9W675;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DE TRANSLANTLY EXPRESSED AXONAL GLYCOPROTEIN.
 GN TAG1.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.
 NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Warren J.T. Jr., Chandrasekhar A., Kanki J.P., Rangarajan R.,
 RA Kuwada J.Y.;
 RT "Molecular cloning and developmental expression analysis of the

RT zebrafish axonal glycoprotein tag1.
 RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF064799; AAD31083.1;
 DR INTERPRO: IPR001777;
 DR INTERPRO: IPR003006;
 DR PFAM: PF00041; In3; 4.
 DR PFAM: PF00047; Ig; 6.
 SQ SEQUENCE 1040 AA: 114626 MW: E6C2DB8B177D1B3C CRC64;

Query Match 71.9%; Score 46; DB 13; Length 1040;
 Best Local Similarity 80.0%; Pred. No. 1.3;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDSGLYQCEA 10
 DB 389 EDSGMQCVN 398

Search completed: June 4, 2001, 12:23:28
 Job time: 591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:31 ; Search time 106.14 Seconds
(without alignments)
6.463 Million cell updates/sec

Title: us-09-284-107-27

Perfect score: 61

Sequence: 1 EDGNNVLRKSPFL 12

Scoring table: BLOSUM62
Gapop 10.0, Capext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : A.Geneseq.0401.*

```

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	12	19	W60559
2	61	100.0	261	20	Y33183
3	61	100.0	344	13	R22549
4	61	100.0	344	17	R91439
5	61	100.0	374	13	R20811
6	61	100.0	374	13	R22550
7	61	100.0	374	17	R91438
8	61	100.0	374	17	W00859
9	61	100.0	374	19	W80448
10	61	100.0	374	19	W97833
11	61	100.0	374	19	W97834

12	61	100.0	374	21	Y96226
13	61	100.0	399	21	B43683
14	53	86.9	344	21	Y96183
15	53	86.9	374	21	Y96134
16	52	85.2	377	20	W86195
17	50	82.0	410	12	R12428
18	42	68.9	149	21	G12850
19	42	68.9	163	21	G12849
20	42	68.9	407	21	G28825
21	42	68.9	421	21	G28824
22	38	62.3	431	21	G13326
23	38	62.3	483	21	G13325
24	38	62.3	523	21	G13324
25	37	60.7	140	21	G58315
26	37	60.7	140	21	G60548
27	37	60.7	261	21	G58314
28	37	60.7	261	21	G60547
29	37	60.7	268	21	G58313
30	37	60.7	268	21	G60546
31	37	60.7	637	21	G52084
32	37	60.7	673	21	G52083
33	37	60.7	715	21	G53159
34	37	60.7	753	21	G53158
35	36	59.0	132	19	Y86103
36	36	59.0	288	17	W14450
37	36	59.0	321	21	G42149
38	36	59.0	358	21	Y94336
39	36	59.0	359	20	Y41690
40	36	59.0	359	21	B44246
41	36	59.0	359	21	B34744
42	36	59.0	507	18	W29772
43	36	59.0	586	21	G53155
44	35	57.4	89	21	G02088
45	35	57.4	156	21	G23166

ALIGNMENTS

```

RESULT 1
W60559 standard; peptide: 12 AA.
W60559:
18-AUG-1998 (first entry)
Oligopeptide from extracellular domain of CD64.
Extracellular domain; CD64; identification; antibody;
immunohistochemical; immunofluorescent analysis; detection;
cell transformation; mutation; anti; oncogene.
Synthetic.
W09815833-A1.
16-APR-1998.
07-OCT-1997; 97WO-NL00557.
08-OCT-1996; 96EP-0202791.
(UYT-) RIKSUNIV UTRCHT.
De Krul CA, Logtenberg T;
WPL: 1998-240964/21.
Identifying peptide(s) binding specifically to protein target - by
expressing on phage surface and testing for binding to immobilised
oligo:peptide derived from the target, useful for, e.g. identifying
specific antibodies

```

XX Example 1; Page 29; 40pp; English.
 PS Synthetic oligopeptides W60537-61 are derived from the extracellular
 CC domain of CD64. They were synthesized on 25 polyethylene rods as
 CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
 CC affinity binding to the CD64-covered rods. The specification describes a
 CC method for the identification of a peptide able to bind specifically to a
 CC target protein. The method comprises displaying the peptide on the
 CC surface of a replicable display package, synthesizing oligopeptides
 CC derived from the target protein on a solid phase, and testing for binding
 CC between the peptide and oligopeptides. The method is used to screen large
 CC peptide libraries, especially to detect antibodies, or their fragments,
 CC that bind to cell markers or that can differentiate between different
 CC forms of the same protein, including bispecific antibodies that bind to
 CC two non-overlapping epitopes on the same monomeric antigen or two
 CC epitopes on different molecules. The genes/oligonucleotides that encode
 CC selected peptides can be isolated and used for recombinant production of
 CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
 CC immunofluorescent analysis, and also to detect cell transformation caused
 CC by mutation in (anti)oncogenes.
 CC
 SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 3.4e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPET 12
 Db 1 edgnvltkrspe1 12

RESULT 2
 Y33183
 ID Y33183 standard; Protein; 261 AA.

XX Y33183;
 DT 15-NOV-1999 (first entry)
 DE Human sFc-gammaRI protein fragment.

XX Fc Receptor; Fc-gammaRIa; human; FcR; model; three-dimension; 3-D;
 KW atomic coordinate; bioactive compound design; computer-assisted;
 KW drug design; therapy; inhibitor; Fc-gamma; Fc-epsilon; IgG; IgE;
 KW tissue damage; hypersensitivity; inflammatory cell recruitment;
 KW inflammatory modulator; Fc-gammaRIa; immune function regulation;
 KW anti-inflammatory; immunoprotective; sFc-gammaRI.
 XX
 OS Homo sapiens.
 XX
 PN WO9940117-A1.

PD 12-AUG-1999.
 PF 04-FEB-1999; 99WO-IB00367.
 PR 11-SEP-1998; 98US-0099994.
 PR 06-FEB-1998; 98US-0073972.
 XX
 PA (ILEX-) ILEXUS PTY LTD.

PI Baell JB, Epa V, Garrett RPJ, Hogarth PM, Matthews BR;
 PI Maxwell KF, McCarthy TD, McKenzie IRC, Pietersz GA;
 PI Powell MS;
 XX

DR WPI; 1999-539978/45.

XX Three-dimensional structures and models of Fc receptors, useful in
 PT computer-assisted drug design
 XX

PS Claim 11; Page 316-317; 326pp; English.

XX This invention describes a novel model of an Fc receptor (FcR) protein
 CC representing a three-dimensional (3-D) structure that substantially
 CC conforms to the specified atomic coordinates. Computer model images of
 CC the FcR can be used to design bioactive chemical compounds, e.g.
 CC oligonucleotides, peptides, peptidomimetics and small organic molecules,
 CC by computer-assisted methods of drug design. Therapeutic compositions
 CC that inhibit the activity of Fc-gammaR or Fc-epsilonR can be used to
 CC reduce IgG- or IgE-mediated tissue damage, e.g. IgG- or IgE-mediated
 CC hypersensitivity, recruitment of inflammatory cells or release of
 CC inflammatory mediators. The therapeutic compositions can also be used to
 CC treat disease in animals. The knowledge of the 3-D structure of Fc-gamma
 CC R1a and models of other FcR provides a means for designing and producing
 CC compounds that regulate immune function and inflammation in an animal,
 CC including humans (i.e. structure based drug design). For example,
 CC chemical compounds can be designed to block binding of immunoglobulin to
 CC an Fc receptor protein using various computer programs and models. The
 CC products of the invention have anti-inflammatory and immunoprotective
 CC activity. This sequence represents the human sFc-gammaRI protein
 CC fragment described in the method of the invention.
 CC
 SQ Sequence 261 AA;

Query Match 100.0%; Score 61; DB 20; Length 261;
 Best Local Similarity 100.0%; Pred. No. 0.001;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPET 12
 Db 246 edgnvltkrspe1 257

RESULT 3
 R22549
 ID R22549 standard; Protein; 344 AA.

XX R22549;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FcRI receptor encoded by clone p98.

XX Rapid immunoselection cloning technique; cell surface antigen;
 KW immunodiagnosis; high affinity receptor.
 KW
 XX

OS Homo sapiens.

PN WO9201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GENO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amlot M;

DR WPI; 1992-056864/07.

DR N-PSDB; Q21179.

PI New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX

PS Example 10; Page 94a; 160pp; English.

XX This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p98/X2. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC while clone p98/X2 predicts a Thr residue. At position 58, p90

CC (see 021180) predicts Val and p98/X2 predicts Leu. Sequences
 CC predicted from all 3 clones show the typical features of a type I
 CC integral membrane protein and include a short hydrophobic signal
 CC sequence, a single 21-residue hydrophobic membrane-spanning domain,
 CC and a short, highly charged cytoplasmic domain. The extracellular
 CC portion contains six potential N-linked glycosylation sites and six
 CC Cys residues distributed among three C2 set Ig-related domains
 CC A fusion protein of FCRI and a receptor ligand will be helpful to
 CC increase the potency of antibodies in therapy.

XX Sequence 344 AA;

Query Match 100.0%; Score 61; DB 13; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPKL 12
 Db 265 edgnvlkrspkl 276

RESULT 4
 ID R91439 standard; Protein; 344 AA.
 AC R91439;
 DT 30-OCT-1996 (first entry)
 XX
 DE Human FCRI (CDNA clone p98 product).
 XX
 KW Cell surface antigen; cloning; immunoselection; immunotherapy;
 KM therapy; diagnosis; vector; FCRI; Fc receptor.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 58
 FT "amino acid 58 is Leu in clone p135
 FT translated product"
 XX
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 23-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI: 1996-200279/20.
 DR N-PSDB: T14718.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX
 PS Example 10; Column 55-56; 79pp; English.
 XX
 CC The amino acid sequence (R91439) of human FCRI was detd. from a
 CC cDNA clone, p98 (T14718), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. 2 other cDNA clones, p135 (T14717) and p90 (T14719),
 CC coded for a variants of the FCRI sequence; the C-terminal sequence
 CC of the p98 product is truncated compared with those of the p135

CC and p90 products (see also R91438 and W00859). A fusion protein
 CC of FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.

XX Sequence 344 AA;

Query Match 100.0%; Score 61; DB 17; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPKL 12
 Db 265 edgnvlkrspkl 276

RESULT 5
 ID R20811 standard; Protein; 374 AA.
 AC R20811;
 DT 21-MAY-1992 (first entry)
 XX
 DE Human macrophage-specific FCRI receptor encoded by clone p135.
 XX
 KW Rapid immunoselection cloning technique; cell surface antigen;
 KM immunodiagnosis; high affinity receptor.
 XX
 OS Homo sapiens.
 XX
 PN W09201049-A.
 PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90WO-US04986.
 PR 13-JUL-1990; 90US-0553759.
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GENO-) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Amlot M;
 XX
 DR WPI: 1992-056864/07;
 DR N-PSDB: Q21178.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of hematopoietic neoplasms, etc.
 XX
 PS Example 10; Page 94a; 160pp; English.
 XX
 CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p135. It differs from the sequence predicted from
 CC clones p98/X2 and p90 (see 021179 and 021180, respectively) at
 CC position 25; p135 encodes a Ser residue and the other two clones
 CC predict a Thr residue. At position 58, p135 predicts Leu and p90
 CC predicts Val. Sequences predicted from all 3 clones show the
 CC typical features of a type I integral membrane protein and include
 CC a short hydrophobic signal sequence, a single 21-residue
 CC hydrophobic membrane-spanning domain, and a short, highly charged
 CC cytoplasmic domain. The extracellular portion contains six
 CC potential N-linked glycosylation sites and six Cys residues
 CC distributed among three C2 set Ig-related domains. A fusion protein
 CC of FCRI and a receptor ligand will be helpful to increase the
 CC potency of antibodies in therapy.

XX Sequence 374 AA;

Query Match 100.0%; Score 61; DB 13; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRSPSEL 12
 DB 265 edgnvklrpsel 276

RESULT 6

ID R22550 standard; Protein: 374 AA

AC R22550;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FCRI receptor encoded by clone p90.

KW Rapid immunoselection cloning technique; cell surface antigen; immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN W09201049-A.

PD 23-JAN-1992.

PF 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PA (GEMO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amiot M;

DR WPI: 1992-056864/07.

DR N-PSDB; Q21180.

PT New CD53 cell surface antigen and DNA encoding it - for immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p90. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC and clone p90 predicts a Thr residue. At position 58, p90 predicts
 CC Val while clones p135 and p98/X2 (see Q21179) both predict Leu.
 CC Sequences predicted from all 3 clones show the typical features of
 CC a type I integral membrane protein and include a short hydrophobic
 CC signal sequence, a single 21-residue hydrophobic membrane-spanning
 CC domain, and a short, highly charged cytoplasmic domain. The
 CC extracellular portion contains six potential N-linked glycosylation
 CC sites and six Cys residues distributed among three C2 set Ig-related
 CC domains. A fusion protein of FCRI and a receptor ligand will be
 CC helpful to increase the potency of antibodies in therapy.

SO Sequence 374 AA;

Query Match 100.0%; Score 61; DB 13; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.0016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRSPSEL 12

DB 265 edgnvklrpsel 276

RESULT 7

ID R91438 standard; Protein: 374 AA.

AC R91438;

DT 30-OCT-1996 (first entry)

XX Human FCRI (cDNA clone p135 product).

KW Cell surface antigen; cloning; immunoselection; immunotherapy; therapy; diagnosis; vector; FCRI; Fc receptor.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Misc-difference 25 /note= "amino acid 25 is Thr in clone p90 and p98 translated products"

FT Misc-difference 58 /note= "amino acid 58 is Val in p90 clone translated product"

PN US5506126-A.

PD 09-APR-1996.

PE 25-FEB-1988; 88US-0160416.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 18-OCT-1993; 93US-0139273.

PA (GEMO-) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

DR WPI: 1996-200279/20.

DR N-PSDB; T14717.

PT Cloning of cDNA encoding cell surface antigen - useful for isolation of diagnostic and therapeutic proteins

PS Example 10; Column 55-56; 79pp; English.

CC The amino acid sequence (R91438) of human FCRI was detd. from a
 CC cDNA clone, p135 (T14717), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the Fc portion of IgG, normally located on cell surfaces of
 CC macrophages. Another isolated cDNA clone, p90 (T14719), coded
 CC for a variant (W00859) of the FCRI sequence, and a third clone, p98
 CC (T14718), coded for an FCRI (R91439) having a different C-terminal
 CC sequence. A fusion protein of FCRI and a receptor ligand will
 CC be helpful in increasing the potency of antibodies in therapy.

SO Sequence 374 AA;

Query Match 100.0%; Score 61; DB 17; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.0016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGNVLRSPSEL 12

DB 265 edgnvklrpsel 276

RESULT 8

ID W00859 standard; Protein: 374 AA.

AC W00859;

DT 30-OCT-1996 (first entry)

DE Human FCRI (cDNA clone p90 product).

KW Cell surface antigen; cloning; immunoselection; immunotherapy;

therapy; diagnosis; vector; FCRI; FC receptor.
 XX Homo sapiens.
 OS US506126-A.
 PN US506126-A.
 PD 09-APR-1996.
 XX 25-FEB-1988; 8805-0160416.
 XX 01-DEC-1992; 9205-0983647.
 PR 25-FEB-1988; 8805-0160416.
 PR 13-JUL-1989; 8905-0379076.
 PR 13-JUL-1990; 9005-0553759.
 PR 18-OCT-1993; 9305-0139273.
 XX (GEMO) GEN HOSPITAL CORP.
 PA Aruffo A. Seed B;
 PI WPI; 1996-200279/20.
 DR N-PSDB; T14719.
 XX Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX Example 10; Column 55-56; 79pp; English.
 PS The amino acid sequence (W00859) of human FCRI was detd. from a
 CC cDNA clone, p90 (T14719), obtd. from a cDNA library using an
 CC immunoselection cloning method. FCRI is a high affinity receptor
 CC for the FC portion of IgG, normally located on cell surfaces of
 CC macrophages. Other cDNA clones (see also T14717-18) coded for
 CC variants (R91438-39) of the FCRI sequence. A fusion protein of
 CC FCRI and a receptor ligand will be helpful in increasing the
 CC potency of antibodies in therapy.
 CC
 SQ Sequence 374 AA;

Query Match 100.0%; Score 61; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0 0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRKSPEL 12
 Db 265 edgnvlkrspel 276

RESULT 9
 W80448 ID W80448 standard; Protein; 374 AA.
 AC W80448;
 XX 07-JUN-1999 (first entry)
 DE Human FC receptor I.
 XX FC receptor I; FCRI; cell surface antigen; lymphocyte; human;
 KW cloning.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT Misc-difference 2 /note- "encoded by TGG"
 FT Misc-difference 23 /note- "encoded by CTC"
 FT Misc-difference 44 /note- "encoded by CTC"
 FT Misc-difference 45 /note- "encoded by GAC"
 FT Misc-difference 45 /note- "encoded by CTG"

Misc-difference 60 /note- "encoded by CCC"
 FT Misc-difference 77 /note- "encoded by AAT"
 FT Misc-difference 85 /note- "encoded by TCC"
 FT Misc-difference 99 /note- "encoded by CAA"
 FT Misc-difference 103 /note- "encoded by CCC"
 FT Misc-difference 141 /note- "encoded by GGC"
 FT Misc-difference 159 /note- "encoded by AAC"
 FT Misc-difference 171 /note- "encoded by ATG"
 FT Misc-difference 176 /note- "encoded by GTC"
 FT Misc-difference 256 /note- "encoded by GGG"
 XX US5830731-A.
 XX 03-NOV-1998.
 PD 21-MAY-1997; 9705-0861205.
 XX 01-DEC-1992; 9205-0983647.
 PR 25-FEB-1988; 8805-0160416.
 PR 13-JUL-1989; 8905-0379076.
 PR 13-JUL-1990; 9005-0553759.
 PR 21-MAY-1997; 9705-0861205.
 XX (GEMO) GEN HOSPITAL CORP.
 PA Aruffo A. Seed B;
 PI WPI; 1998-609251/51.
 DR N-PSDB; V63456.
 XX New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNAs, especially
 PT human lymphocyte antigenic sequences
 XX Example 10; Column 53-54; 75pp; English.
 PS This is the amino acid sequence of human FC receptor I (FCRI), as
 CC deduced from cDNA clone p135 (see V63456) isolated using a rapid
 CC immunoselection cloning method from a cDNA library expressed in COS
 CC cells. The cDNA library was constructed from polyA RNA of cells
 CC from a single patient undergoing extracorporeal Interleukin-2
 CC induction therapy. Clones p90 (see X07372) and p98 (see X07373)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed, and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 XX Sequence 374 AA;

Query Match 100.0%; Score 61; DB 19; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRSPSEL 12
DB 265 edgnvltkrspe1 276

RESULT 10
ID W97833 standard; Protein: 374 AA.

XX W97833;

XX 07-JUN-1999 (first entry)

DE Human Fc receptor I.

XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
cloning.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 2 /note= "encoded by TGG"

FT Misc-difference 23 /note= "encoded by CTC"

FT Misc-difference 44 /note= "encoded by GAC"

FT Misc-difference 45 /note= "encoded by CTG"

FT Misc-difference 60 /note= "encoded by CCC"

FT Misc-difference 77 /note= "encoded by AAT"

FT Misc-difference 85 /note= "encoded by TCC"

FT Misc-difference 99 /note= "encoded by CAA"

FT Misc-difference 103 /note= "encoded by CCC"

FT Misc-difference 141 /note= "encoded by GGC"

FT Misc-difference 159 /note= "encoded by AAC"

FT Misc-difference 171 /note= "encoded by ATG"

FT Misc-difference 176 /note= "encoded by GTC"

FT Misc-difference 256 /note= "encoded by GGC"

PN US5830731-A.

XX 03-NOV-1998.

XX 21-MAY-1997; 97US-0861205.

XX 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

XX 13-JUL-1989; 89US-0379076.

XX 13-JUL-1990; 90US-0553759.

XX 21-MAY-1997; 97US-0861205.

XX (GEHO) GEN HOSPITAL CORP.

XX Aruffo A, Seed B;

XX WPI; 1998-609251/51.

XX N-PSDB; X07372.

New cloning vector and polylinker - based on existing sequences for

PT efficient cloning and expression of mammalian cDNA(s), especially
PT human lymphocyte antigenic sequences

Example 10; Column 53-54; 75pp; English.

This is the amino acid sequence of human Fc receptor I (FCRI), as deduced from cDNA clone p90 (see X07372) isolated using a rapid immunoselection cloning method from a cDNA library expressed in COS cells. The cDNA library was constructed from polyA RNA of cells from a single patient undergoing extracorporeal interleukin-2 induction therapy. Clones p135 (see V63456) and p98 (see X07373) were also obtained. A novel method for cloning cDNAs from mammalian expression libraries is based on transient expression of an antigen in eukaryotic cells and selection of cells expressing the antigen by adhesion to an antibody-coated substrate. The method is useful for the isolation and cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell. It has been used to clone genes (see V63442-63) encoding cell surface antigens from mammalian lymphocytes (see W80440-55). The isolated genes can be expressed in a prokaryotic or eukaryotic host cells to produce the encoded protein. The invention also provides high efficiency expression vectors (see V63441 and V63444) which allow the generation of very large mammalian expression libraries. The purified genes and proteins are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders of animals, including humans.

Sequence 374 AA:

Query Match 100.0%; Score 61; DB 19; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.0016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRSPSEL 12
DB 265 edgnvltkrspe1 276

RESULT 11

ID W97834 standard; Protein: 374 AA.

XX W97834;

XX 07-JUN-1999 (first entry)

DE Human Fc receptor I.

XX Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;
cloning.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 2 /note= "encoded by TGG"

FT Misc-difference 23 /note= "encoded by CTC"

FT Misc-difference 44 /note= "encoded by GAC"

FT Misc-difference 45 /note= "encoded by CTG"

FT Misc-difference 60 /note= "encoded by CCC"

FT Misc-difference 77 /note= "encoded by AAT"

FT Misc-difference 85 /note= "encoded by TCC"

FT Misc-difference 99 /note= "encoded by CAA"

FT Misc-difference 103

```

FT      /note= "encoded by CCC"
FT      Misc-difference 141
FT      /note= "encoded by GGC"
FT      Misc-difference 159
FT      /note= "encoded by AAC"
FT      Misc-difference 171
FT      /note= "encoded by ATG"
FT      Misc-difference 176
FT      /note= "encoded by CTC"
FT      Misc-difference 256
FT      /note= "encoded by GCG"
FT      US830731-A.
PN
XX
XX      03-NOV-1998.
PD
XX
XX      21-MAY-1997; 97US-0861205.
PE
XX
XX      01-DEC-1992; 92US-0983647.
PR      25-FEB-1988; 88US-0160416.
PR      13-JUL-1989; 89US-0379076.
PR      13-JUL-1990; 90US-0553759.
PR      21-MAY-1997; 97US-0861205.
XX
XX      (GEHO ) GEN HOSPITAL CORP.
XX
XX      Arnolfo A, Seed B;
PI
XX      WPI; 1998-609251/51.
DR      N-PSDB; X07373.
XX
XX      New cloning vector and polylinker - based on existing sequences for
PT      efficient cloning and expression of mammalian cDNA(s), especially
PT      human lymphocyte antigenic sequences
XX
XX      Example 10; Column 53-54; 75pp; English.
PS
XX
XX      This is the amino acid sequence of human Fc receptor I (FcRI), as
CC      deduced from cDNA clone p88/X2 (see X07373) isolated using a rapid
CC      immunoselection cloning method from a cDNA library expressed in COS
CC      cells. The cDNA library was constructed from polyA RNA of cells
CC      from a single patient undergoing extracorporeal interleukin-2
CC      induction therapy. Clones p135 (see V63456) and p90 (see X07372)
CC      were also obtained. A novel method for cloning cDNAs from mammalian
CC      expression libraries is based on transient expression of an antigen
CC      in eukaryotic cells and selection of cells expressing the antigen by
CC      adhesion to an antibody-coated substrate. The method is useful for
CC      the isolation and cloning of any protein which can be expressed and
CC      transported to the cell surface membrane of a eukaryotic cell. It
CC      has been used to clone genes (see V63442-63) encoding cell surface
CC      antigens from mammalian lymphocytes (see W80440-53). The isolated
CC      genes can be expressed in a prokaryotic or eukaryotic host cells to
CC      produce the encoded protein. The invention also provides high
CC      efficiency expression vectors (see V63441 and V63444) which allow
CC      the generation of very large mammalian expression libraries. The
CC      purified genes and proteins are useful for immunodiagnostic and
CC      immunotherapeutic applications, including the diagnosis and
CC      treatment of immune-mediated infections, diseases, and disorders of
CC      animals, including humans.
XX
XX      Sequence 374 AA;
SQ

```

```

Query Match      100.0%; Score 61; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 EDGNVLRSPSL 12
      |||
DB      265 edgnvlkrspsl 276

```

RESULT 12

```

Y96226
ID      Y96226 standard; Protein; 374 AA.
XX
XX      Y96226;
AC
XX
XX      11-SEP-2000 (first entry)
DT
XX
XX      Human high affinity Fc receptor, FcgammaRI.
DE
XX
XX      Human; high affinity Fc receptor; FcgammaRI; immunoglobulin;
KW      infection; immune response; CD64; monocyte; macrophage; neutrophil;
KW      eosinophil; HIV; IgG; immunosuppressive; antirheumatic; cytostatic;
KW      antineoplastic; AIDS; SLE; multiple myeloma; rheumatoid arthritis;
KW      systemic lupus erythematosus; tumour.
XX
XX      Homo sapiens.
OS
XX
XX      EPI006183-A1.
PN
XX
XX      07-JUN-2000.
PD
XX
XX      03-DEC-1998; 98EP-0122969.
PE
XX
XX      03-DEC-1998; 98EP-0122969.
PR
XX
XX      (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX      WPI; 2000-367968/32.
DR      N-PSDB; A27466.
XX
XX      Novel Fc receptor lacking transmembrane domains, a signal peptide, and
PT      glycosylation, useful for diagnosing and treating immune disorders and
PT      cancer
XX
XX      Disclosure; Page 26-28; 60pp; English.
PS
XX
XX      The present sequence is the human high affinity Fc receptor, FcgammaRI.
CC      FcgammaRI is also known as CD64. Fc receptors play an important
CC      role in defending the body against infections. First, pathogens are
CC      opsonised by serum immunoglobulins. The resulting complex then binds to
CC      cells expressing Fc receptors. FcgammaRI molecules are expressed by
CC      monocytes and macrophages, but expression can also be induced on
CC      neutrophils and eosinophils. Upon Fc receptor activation, immune effector
CC      pathways are activated, leading to immune response. The present sequence
CC      may be modified to produce recombinant versions. The recombinant Fc
CC      receptor consist only of the extracellular portion of the receptor and
CC      are not glycosylated i.e. they do not have transmembrane domains or
CC      signal peptides. The recombinant proteins may be used in immunoassays to
CC      determine the immune status of patients with chronic diseases of the
CC      immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple
CC      myeloma (MM), or rheumatoid arthritis. In addition, pharmaceutical
CC      compositions containing recombinant proteins may be used to treat or
CC      prevent autoimmune diseases, allergies or tumours, especially AIDS,
CC      rheumatoid arthritis or MM.
XX
XX      Sequence 374 AA;
SQ

```

```

Query Match      100.0%; Score 61; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 EDGNVLRSPSL 12
      |||
DB      265 edgnvlkrspsl 276

```

```

RESULT 13
B43683
ID      B43683 standard; Protein; 399 AA.
XX
XX      B43683;
AC
XX

```

DT 08-FEB-2001 (first entry)
 XX Human cancer associated protein sequence SEQ ID NO:1128.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antidiabetic; antirheumatic; antidiabetic; antiviral;
 KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;
 KW vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 XX WO200005350-A1.
 XX
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05882.
 XX
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX
 XX WPI: 2000-587533/55.
 XX
 XX N-PSDB: C77892.
 XX
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 XX useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 11; Page 1739-1740; 2352pp; English
 XX
 CC C77607 to C78448 encode the human cancer associated proteins given in
 CC B43396 to B44239. The proteins can have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC cytostatic; proliferative; vulnery; immunomodulator; antidiabetic;
 CC antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 CC antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 CC neuroprotective; cardiac; thrombolytic; coagulant; nocotropic;
 CC vasotropic; antiproliferative and angiogenic. The polynucleotides and
 CC polypeptides can be used for preventing, treating or ameliorating medical
 CC conditions and diagnosis pathological conditions. Polynucleotides,
 CC polypeptides, antibodies, agonists and antagonists from the present
 CC invention may be used to treat immune disorders by activating or
 CC inhibiting the proliferation, differentiation or mobilisation of immune
 CC cells, to treat disorders of haematopoietic cells, autoimmune disorders,
 CC allergic reactions, graft versus host disease and organ rejection,
 CC modulate haemostatic or thrombolytic activity, modulate inflammation,
 CC cancers, cardiovascular disorders, neurological disease and bacterial or
 CC viral infections. The peptides, nucleotides, antibodies, agonists and
 CC antagonists may be also be used in drug screens. C78449 to C78457 and
 CC B44240 represent sequences used in the exemplification of the present
 CC invention.
 CC
 XX
 XX Sequence 399 AA;
 SQ

Query Match 100.0%; Score 61; DB 21; Length 399;
 Best Local Similarity 100.0%; Fred. No. 0.0017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRSPFL 12
 |||||
 Db 290 edgnvkrspel 301

RESULT 14
 Y96183

ID Y96183 standard; Protein; 344 AA.
 XX
 AC Y96183;
 XX
 DT 19-DEC-2000 (first entry)
 XX
 DE Human macrophage-specific FCRI.
 KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
 KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; parasitic disease; FC receptor.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH 15
 FT Misc-difference 15
 FT /note= "encoded by GCG"
 FT
 FT Misc-difference 38
 FT /note= "encoded by ACC"
 FT
 FT Misc-difference 50
 FT /note= "encoded by CCG"
 FT
 FT Misc-difference 51
 FT /note= "encoded by ACC"
 FT
 FT Misc-difference 55
 FT /note= "encoded by CAC"
 FT
 FT Misc-difference 56
 FT /note= "encoded by TCC"
 FT
 FT Misc-difference 60
 FT /note= "encoded by CCC"
 FT
 FT Misc-difference 64
 FT /note= "encoded by CAG"
 FT
 FT Misc-difference 82
 FT /note= "encoded by CAA"
 FT
 FT Misc-difference 116
 FT /note= "encoded by GAA"
 FT
 FT Misc-difference 117
 FT /note= "encoded by GCA"
 FT
 FT Misc-difference 122
 FT /note= "encoded by TTC"
 FT
 FT Misc-difference 123
 FT /note= "encoded by ACC"
 FT
 FT Misc-difference 126
 FT /note= "encoded by CCG"
 FT
 FT Misc-difference 129
 FT /note= "encoded by CAG"
 FT
 FT Misc-difference 134
 FT /note= "encoded by AAT"
 FT
 FT Misc-difference 136
 FT /note= "encoded by GTT"
 FT
 FT Misc-difference 139
 FT /note= "encoded by CCA"
 FT
 FT Misc-difference 140
 FT /note= "encoded by AAT"
 FT
 FT Misc-difference 213
 FT /note= "encoded by CAA"
 FT
 FT Misc-difference 216
 FT /note= "encoded by TTC"
 FT
 FT Misc-difference 220
 FT /note= "encoded by CGT"
 FT
 FT Misc-difference 268
 FT /note= "encoded by AAT"
 FT
 FT Misc-difference 305
 FT /note= "encoded by GTG"
 FT
 FT Misc-difference 306
 FT /note= "encoded by AAC"
 FT
 FT Misc-difference 332
 FT /note= "encoded by GGT"
 FT
 FT Misc-difference 333
 FT /note= "encoded by GGC"
 FT
 FT Misc-difference 338
 FT /note= "encoded by CCT"
 XX

OS	Homo sapiens.
XX	Location/Qualifiers
FH	Misc-difference 15 /note= "encoded by GCG"
FT	Misc-difference 38 /note= "encoded by ACC
FT	Misc-difference 50 /note= "encoded by CCG"
FT	Misc-difference 51 /note= "encoded by ACC
FT	Misc-difference 55 /note= "encoded by CAC
FT	Misc-difference 56 /note= "encoded by TTC"
FT	Misc-difference 60 /note= "encoded by CCC
FT	Misc-difference 64 /note= "encoded by CAG
FT	Misc-difference 82 /note= "encoded by CAA"
FT	Misc-difference 116 /note= "encoded by GAA
FT	Misc-difference 117 /note= "encoded by GCA"
FT	Misc-difference 122 /note= "encoded by TTT"
FT	Misc-difference 123 /note= "encoded by ACC
FT	Misc-difference 126 /note= "encoded by CCG"
FT	Misc-difference 129 /note= "encoded by CAT"
FT	Misc-difference 134 /note= "encoded by AAT"
FT	Misc-difference 136 /note= "encoded by GTT"
FT	Misc-difference 139 /note= "encoded by CCA"
FT	Misc-difference 140 /note= "encoded by AAT"
FT	Misc-difference 213 /note= "encoded by CAA"
FT	Misc-difference 216 /note= "encoded by TTC
FT	Misc-difference 220 /note= "encoded by CGT"
FT	Misc-difference 268 /note= "encoded by AAT"
FT	Misc-difference 305 /note= "encoded by GTG
FT	Misc-difference 306 /note= "encoded by AAC
FT	Misc-difference 332 /note= "encoded by GGT"
PN	US611093-A.
XX	
PD	29-AUG-2000.
PF	
PE	28-OCT-1998; 98U5-0181612.
PR	
PR	01-DEC-1992; 92U5-0983647. 88U5-0160416. 25-FEB-1988; 13-JUL-1989; 89O5-0379076. 23-MAR-1990; 90O5-0498809. 13-JUL-1990; 90O5-0553759.
PA	(GEHO) GEN HOSPITAL CORP.
XI	-Stamenkovic I, Seed B;
WP1;2000-586382/55.	

DR N-PSDB; A50592.

XX
XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
PT useful for immunodiagnosis and immunotherapy of immune-mediated
PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
PT diseases.

XX
XX
PS Example 10; Column 53-55; 75pp; English.

CC The present sequence is that of a human macrophage specific FCRI,
CC as deduced from cDNA clone p135 (see A50592), which was isolated
CC from a cDNA library expressed in COS cells using a novel method of
CC the invention designed to isolate CSA nucleic acids. The method is
CC based upon transient expression of a CSA in eukaryotic cells and
CC physical selection of cells expressing the antigen by adhesion to
CC (panning on) an antibody-coated substrate such as a culture dish.
CC CSA nucleic acids isolated by the method of the invention, and the
CC proteins they encode, are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders in
CC animals, including humans. These disorders include asthma, multiple
CC immune-complex disease, amyloidosis, parasitic diseases or multiple
CC IgG, normally located on the cell surfaces of macrophages. The
CC ability to interfere with such bonding, or to cause it to occur on
CC surfaces other than macrophages, is useful in therapy. A fusion
CC protein of FCRI and a receptor ligand will be helpful to increase
CC the potencies of antibodies in therapy.

XX
XX
S0 Sequence 374 AA:

Query Match

Best Local Similarity 86.9%; Score 53; DB 21; Length 374;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EDGNVTKRSPSL 12

Db 265 EDGNVTKRSPSL 276

Search completed: June 4, 2001, 12:13:32
Job time: 205 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:31 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-27

Perfect score: 61

Sequence: 1 EDGNVIAKRSPDL 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfilltest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	82.0	91	2	US-08-332-562A-94
2	36	59.0	213	2	US-08-531-525-11
3	36	59.0	213	2	US-08-718-270A-11
4	36	59.0	288	2	US-08-737-825-8
5	35	57.4	479	2	US-08-394-189B-6
6	35	57.4	497	3	US-08-258-287B-37
7	35	57.4	497	3	US-08-368-704C-37
8	35	57.4	503	3	US-08-258-287B-36
9	35	57.4	503	3	US-08-368-704C-36
10	35	57.4	505	2	US-08-394-189B-5
11	35	57.4	505	5	PCT-US93-05701-20
12	35	57.4	505	5	PCT-US93-05705-5
13	35	57.4	1865	1	US-08-588-985-2
14	35	57.4	1865	1	US-08-971-988-2
15	35	57.4	2710	2	US-08-568-459A-12
16	35	57.4	2710	2	US-08-487-826B-12
17	35	57.4	3060	2	US-08-487-826B-14
18	34	55.7	503	2	US-08-781-802-2
19	34	55.7	503	2	US-08-781-802-10
20	34	55.7	503	2	US-08-781-802-12
21	34	55.7	566	3	US-08-926-842B-18
22	34	55.7	566	3	US-08-926-842B-19
23	33	54.1	318	2	US-08-872-719-2
24	33	54.1	318	2	US-08-957-302A-12
25	33	54.1	319	2	US-09-336-890-2
26	33	54.1	319	2	US-08-795-927-4
27	33	54.1	480	4	US-09-182-859-4

28	33	54.1	480	4	US-09-170-670-5	Sequence 5, Appl1
29	33	54.1	480	4	US-09-193-068-5	Sequence 5, Appl1
30	33	54.1	480	4	US-09-183-412-5	Sequence 5, Appl1
31	33	54.1	489	5	PCT-US96-10521-9	Sequence 9, Appl1
32	33	54.1	503	3	US-08-258-287B-2	Sequence 2, Appl1
33	33	54.1	503	3	US-08-258-287B-35	Sequence 35, Appl1
34	33	54.1	503	3	US-08-368-704C-2	Sequence 2, Appl1
35	33	54.1	503	3	US-08-368-704C-35	Sequence 35, Appl1
36	33	54.1	503	5	PCT-US93-05701-19	Sequence 19, Appl1
37	33	54.1	503	5	PCT-US93-05705-2	Sequence 2, Appl1
38	33	54.1	514	1	US-08-720-899-4	Sequence 4, Appl1
39	33	54.1	514	1	US-08-459-610-4	Sequence 4, Appl1
40	33	54.1	514	2	US-08-343-804-4	Sequence 4, Appl1
41	33	54.1	514	2	US-08-687-399-4	Sequence 4, Appl1
42	33	54.1	514	2	US-08-600-908A-4	Sequence 4, Appl1
43	33	54.1	514	3	US-08-683-838A-4	Sequence 4, Appl1
44	33	54.1	520	1	US-08-468-700-36	Sequence 36, Appl1
45	33	54.1	520	5	PCT-US94-01553A-34	Sequence 34, Appl1

ALIGNMENTS

RESULT 1
US-08-332-562A-94
Sequence 94, Application US/08332562A
Patent No. 5985599
GENERAL INFORMATION:
APPLICANT: MCKENZIE, Ian F.C.
APPLICANT: HOGARTH, Mark P.
APPLICANT: HIRBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HOLETT, Mark D.
TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-94
Query Match 82.0%, Score 50, DB 2, Length 91;

Best Local Similarity 83.3%; Pred. No. 0.0086;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGNVLRKSP 12
DB 77 EDSSVLRKSP 88

RESULT 2

US-08-531-525-11

Sequence 11, Application US/08531525

Patent No. 5840683

GENERAL INFORMATION:

APPLICANT: Hlavka, Joseph J.

APPLICANT: Pincus, Matthew R.

APPLICANT: No. 5840683le, John F.

APPLICANT: Abajian, Henry B.

APPLICANT: Kende, Andrew S.

TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action

TITLE OF INVENTION: of P21 Ras

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee and Winner, P.C.

STREET: 5370 Manhattan Circle, Suite 201

CITY: Boulder

STATE: Colorado

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/531,525

FILING DATE: 21-SEP-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Feider, Donna M.

REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 37-94

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 213 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Arabidopsis thaliana

US-08-531-525-11

Query Match 59.0%; Score 36; DB 2; Length 213;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRK 8

DB 200 EDGSVLRK 207

RESULT 3

US-08-718-270A-11

Sequence 11, Application US/08718270A

Patent No. 5910478

GENERAL INFORMATION:

APPLICANT: Hlavka, Joseph J.

APPLICANT: Pincus, Matthew R.

APPLICANT: No. 5910478le, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptidomimetics Inhibiting
TITLE OF INVENTION: the Oncogenic Action of P21 Ras
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,270A
FILING DATE: 20-SEP-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/531,525
FILING DATE: 21-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,091
FILING DATE: 21-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Feider, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 78-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 213 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana

US-08-718-270A-11

Query Match 59.0%; Score 36; DB 2; Length 213;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRK 8

DB 200 EDGSVLRK 207

RESULT 4

US-08-737-825-8

Sequence 8, Application US/08737825

Patent No. 5871922

GENERAL INFORMATION:

APPLICANT: SALMOND, GEORGE PEACOCK COPELAND

APPLICANT: MCGOWAN, SIMON JAMES

APPLICANT: SEBATHIA, MOHAMED

APPLICANT: COX, ANTHONY RICHARD JOHN

APPLICANT: HOLDEN, MATTHEW THOMAS GEOFFREY

APPLICANT: PORTER, LAUREN ELIZABETH

APPLICANT: BYCROFT, BARIE WALSHAM

APPLICANT: WILLIAMS, PAUL

APPLICANT: STEWART, GORDON SIDNEY ANDERSON BIRNIE

TITLE OF INVENTION: GENES INVOLVED IN THE BIOSYNTHETIC PATHWAY OF CARBAPENEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD97
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/737,825
FILING DATE: 03-JUN-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 1009-0105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE: NO
ORGANISM: Erwinia carotovora

US-08-737-825-8

Query Match 59.0%; Score 36; DB 2; Length 288;

Best Local Similarity 70.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DGNVLKRSPE 11

DB 137 DGNVLKRSPE 146

RESULT 5

US-08-394-189B-6

Sequence 6, Application US/08394189B

Patent No. 5962301

GENERAL INFORMATION:

APPLICANT: Horvitz, Robert

APPLICANT: Yuan, Junying

APPLICANT: Shiham, Shai

TITLE OF INVENTION: RELATDNSS OF HUMAN INTERLEUKIN-1

TITLE OF INVENTION: BETA CONVERTASE GENE TO A C. ELLEBANS CELL DEATH

TITLE OF INVENTION: GENE, INHIBITORY PORTIONS OF THESE GENES AND...

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/394,189B

FILING DATE: 24-FEB-2005

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/282,211

FILING DATE: 12-JUL-1994

APPLICATION NUMBER: 07/984,182

FILING DATE: 20-NOV-1992

APPLICATION NUMBER: 07/697,788

FILING DATE: 12-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Bleker-Brady, Kristina

REGISTRATION NUMBER: 39,109

REFERENCE/DOCKET NUMBER: 01997/211001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX: 617-428-7045

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 479 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-394-189B-6

Query Match 57.4%; Score 35; DB 2; Length 479;

Best Local Similarity 50.0%; Pred. No. 53;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDGNVLKRSPEL 12

DB 447 QGANILKMPPEL 458

RESULT 6

US-08-258-287B-37

Sequence 37, Application US/08258287B

Patent No. 6083735

GENERAL INFORMATION:

APPLICANT: Yuan, Junying

APPLICANT: Miura, Masayuki

TITLE OF INVENTION: Programmed Cell Death Genes and Proteins

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/258,287B

FILING DATE: 10-JUN-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/080,850

FILING DATE: 24-JUN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Bugalsky, Lawrence B.

REGISTRATION NUMBER: 35,086

REFERENCE/DOCKET NUMBER: 0609.3920001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

TELEX: 248636 SSK

INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: both
US-08-258-287B-37

Query Match 57.4%; Score 35; DB 3; Length 497;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPDL 12
DB 465 OGNNILKQMPDL 476

RESULT 7
US-08-368-704C-37
Sequence 37, Application US/08368704C
Patent No. 6087160
GENERAL INFORMATION:
APPLICANT: Yuan, Junying
APPLICANT: Mura, Masayuki
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,704C
FILING DATE: 4-JAN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,287
FILING DATE: 10-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/080,850
FILING DATE: 24-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609.3920002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: both

US-08-368-704C-37

Query Match 57.4%; Score 35; DB 3; Length 497;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPDL 12
DB 465 OGNNILKQMPDL 476

RESULT 8

US-08-258-287B-36
Sequence 36, Application US/08258287B
Patent No. 6083735
GENERAL INFORMATION:
APPLICANT: Yuan, Junying
APPLICANT: Mura, Masayuki
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,287B
FILING DATE: 10-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/080,850
FILING DATE: 24-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609.3920001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: both

US-08-258-287B-36

Query Match 57.4%; Score 35; DB 3; Length 503;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPDL 12
DB 471 OGNNILKQMPDL 482

RESULT 9
US-08-368-704C-36

Sequence 36, Application US/08368704C
Patent No. 6087160
GENERAL INFORMATION:
APPLICANT: Yuan, Junying
APPLICANT: Mura, Masayuki
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,704C
FILING DATE: 4-JAN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,287
FILING DATE: 10-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/080,850
FILING DATE: 24-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: both
US-08-368-704C-36

Query Match 57.4%; Score 35; DB 3; Length 503;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPPEL 12
Db 471 QGSNLIKOMPPEL 482

RESULT 10
US-08-394-189B-5
Sequence 5, Application US/08394189B
Patent No. 5962301
GENERAL INFORMATION:
APPLICANT: Horvitz, Robert
APPLICANT: Yuan, Junying
APPLICANT: Shaham, Shai
TITLE OF INVENTION: RELATEDNESS OF HUMAN INTERLEUKIN-1
TITLE OF INVENTION: BETA CONVERTASE GENE TO A C. ELEGANS CELL DEATH
TITLE OF INVENTION: GENE, INHIBITORY PORTIONS OF THESE GENES AND...
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elding LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,189B
FILING DATE: 24-FEB-2005
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/282,211
FILING DATE: 12-JUL-1994
APPLICATION NUMBER: 07/984,182
FILING DATE: 20-NOV-1992
APPLICATION NUMBER: 07/897,788
FILING DATE: 12-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bieker-Brady, Kristina

REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 01997/211001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-394-189B-5

Query Match 57.4%; Score 35; DB 2; Length 505;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPPEL 12
Db 473 QGSNLIKOMPPEL 484

RESULT 11
PCT-US93-05701-20
Sequence 20, Application PC/TUS9305701
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: Cloning and Characterization of Cell Death Genes
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Massachusetts Institute of Technology
STREET: 77 Massachusetts Avenue
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER:
OPERATING SYSTEM:
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05701
FILING DATE: 19930614
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: unsure
LOCATION: at every Xaa
PCT-US93-05701-20

Query Match 57.4%; Score 35; DB 5; Length 505;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPPEL 12
Db 473 QGSNLIKOMPPEL 484

RESULT 12
PCT-US93-05705-5
Sequence 5, Application PC/TUS9305705
GENERAL INFORMATION:

APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: Inhibitors of Ccd-3 and Related Proteins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Massachusetts Institute of Technology
STREET: 77 Massachusetts Avenue
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05705
FILING DATE: 1993/07/14
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: unsure
LOCATION: at every Xaa
PCT-US93-05705-5

Query Match 57.4%; Score 35; DB 5; Length 505;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 EDGNVLRSPKL 12
DB 473 QGSNLIKMPKL 484

RESULT 13
US-08-588-985-2
Sequence 2, Application US/08588985
Patent No. 5777094
GENERAL INFORMATION:
APPLICANT: MICHIOYUKI MATSUDA et al.
TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,985
FILING DATE: January 19, 1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 1865 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: spleen cell of homo sapiens
US-08-588-985-2

Query Match 57.4%; Score 35; DB 1; Length 1865;
Best Local Similarity 66.7%; Pred. No. 2,5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;
QY 1 EDGNVLRSPKL 12
DB 169 EDGNIL--DPKL 178

RESULT 14
US-08-971-988-2
Sequence 2, Application US/08971988
Patent No. 5786461
GENERAL INFORMATION:
APPLICANT: MICHIOYUKI MATSUDA et al.
TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,988
FILING DATE: 17-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/588,985
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1865 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: spleen cell of homo sapiens
US-08-971-988-2

Query Match 57.4%; Score 35; DB 1; Length 1865;
Best Local Similarity 66.7%; Pred. No. 2,5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 1 EDGNVLRSPRL 12
1111:1111
Db 169 EDGNIL--DPRL 178

RESULT 15

US-08-568-459A-12
Sequence 12, Application US/08568459A
Patent No. 5849306

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

US-08-568-459A-12

Query Match 57.4%; Score 35; DB 2; Length 2710;

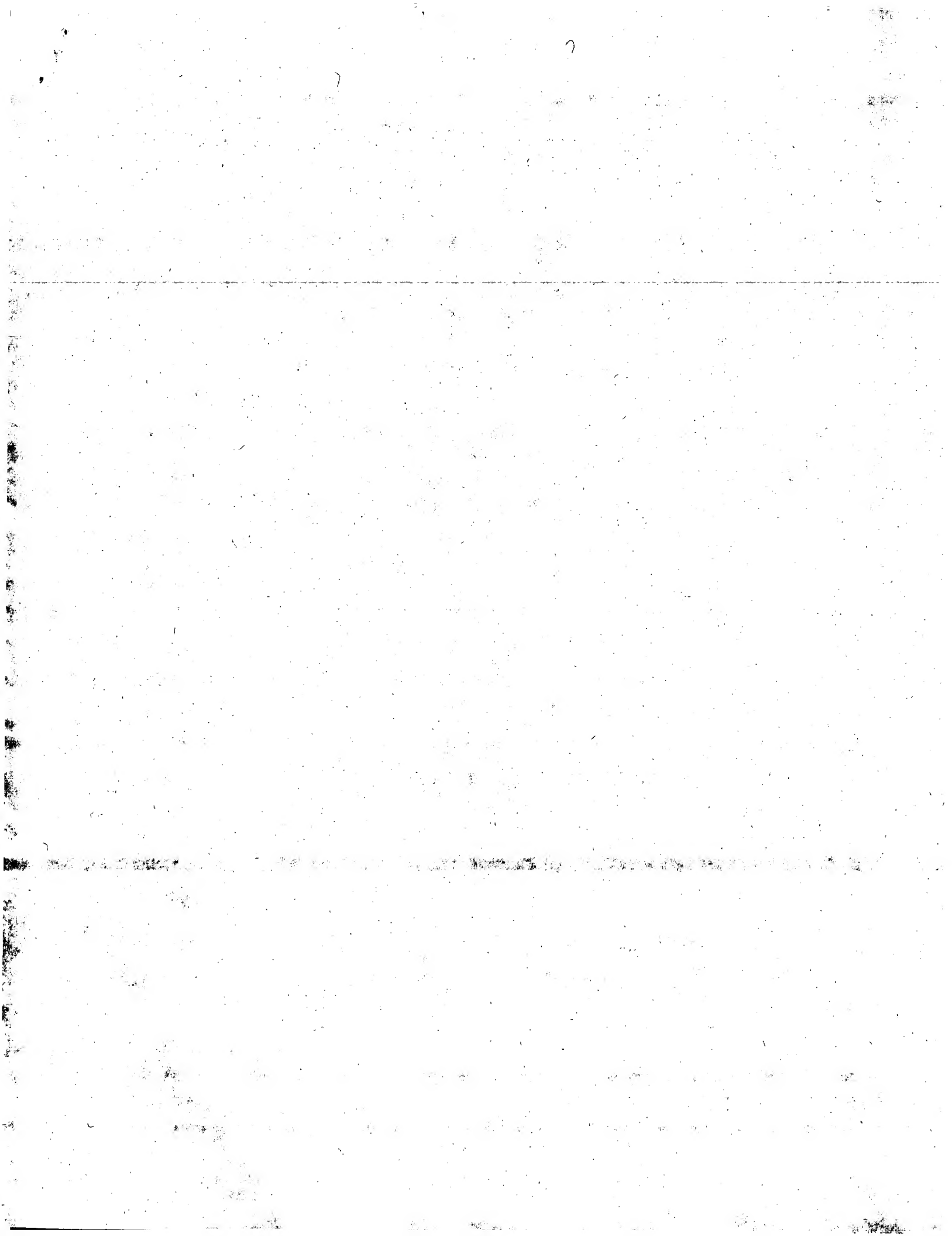
Best Local Similarity 70.0%; Pred. No. 3.9e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DGNVLRSPRL 11
1111:1111

Db 1680 DGNVLRSPRL 1689

Search completed: June 4, 2001, 12:14:32
Job time: 225 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:44 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-27

Sequence: 1 EDGNVLRKSPSL 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.67:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	61	100.0	344	2	A41357	Fc gamma (IgG) rec
2	61	100.0	374	1	A39878	Fc gamma (IgG) rec
3	50	82.0	404	2	A46480	Fc gamma (IgG) rec
4	42	68.9	412	2	S55946	probable membrane
5	42	68.9	407	2	S77922	shaggy-like protei
6	41	67.2	336	2	I48471	Fc gamma (IgG) rec
7	40	65.6	396	2	H82339	conserved hypotnet
8	40	65.6	651	2	D72605	probable long-chain
9	38	62.3	1314	2	T09481	mating type silenc
10	37.5	61.5	146	2	A25233	replication protei
11	37.5	61.5	158	2	A29827	replication protei
12	37.5	61.5	158	2	A61069	replication protei
13	37.5	61.5	162	2	B24497	replication protei
14	37	60.7	181	2	S19534	crustacyanin chain
15	37	60.7	268	2	T05123	hypothetical prote
16	37	60.7	388	1	SYEC5B	succinate--CoA lig
17	37	60.7	871	2	T32275	hypothetical prote
18	37	60.7	1354	2	T13930	tripeptidyl-peptid
19	37	60.7	1970	2	T03284	myoblast city prot
20	36	59.0	154	2	J50710	heat shock protein
21	36	59.0	159	2	S72544	heat shock protein
22	36	59.0	160	2	JC4377	low-molecularweig
23	36	59.0	216	2	J50639	GTP-binding protei
24	36	59.0	276	2	F83578	hypothetical prote
25	36	59.0	290	2	D82545	succinyl-CoA synth
26	36	59.0	687	2	T29220	hypothetical prote
27	36	59.0	902	2	C83964	cation-transportin
28	35.5	58.2	159	2	C71192	probable molybdenu
29	35	57.4	156	2	J00351	heat shock protein

30	35	57.4	156	2	T49264	heat shock protein
31	35	57.4	157	2	S06074	heat shock protein
32	35	57.4	157	2	T14381	heat shock protein
33	35	57.4	159	2	B84096	mutator MucT prote
34	35	57.4	166	2	T48040	hypothetical prote
35	35	57.4	197	2	S19102	hypothetical prote
36	35	57.4	202	2	T36138	hypothetical prote
37	35	57.4	299	2	S25782	hypothetical prote
38	35	57.4	299	2	D23686	lact protein - Lac
39	35	57.4	370	2	C40585	recf protein - Sta
40	35	57.4	376	2	S73941	oligopeptide trans
41	35	57.4	384	2	T00625	branched-chain ami
42	35	57.4	400	2	T34363	hypothetical prote
43	35	57.4	424	2	J00661	lmpb protein - Sal
44	35	57.4	454	2	A46498	glucocorticoid-sen
45	35	57.4	546	2	C72453	hypothetical prote

ALIGNMENTS

```

RESULT 1
A41357
Fc gamma (IgG) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000.
C:Accession: A41357; S03019
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A:Title: Isolation and expression of functional high-affinity Fc receptor complementa
A:Reference number: A41357; MUID:89100284
A:Accession: A41357
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: GB:X14355; GB:M21090; NID:931333; PIDN:CAA32536.1; PID:931334
R:Allen, J.M.; Seed, B. 11824, 1988
Nucleic Acids Res. 16, 11824, 1988
A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (
A:Reference number: S03018; MUID:89098339
A:Accession: S03019
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: EMBL:X14355; NID:931333; PIDN:CAA32536.1; PID:931334
A:Note: the authors translated the codon ACT for residue 25 as Ala
C:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
F:117-170/Domain: immunoglobulin homology <IMW>

Query Match 100.0%; Score 61; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 EDGNVLRKSPSL 12
Db 265 EDGNVLRKSPSL 276

RESULT 2
A39878
Fc gamma (IgG) receptor I-A (high affinity) precursor - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: A39878; I70304; B41357; S03018; I57525
R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
J. Biol. Chem. 266, 13449-13455, 1991
A:Title: Gene organization of the human high affinity receptor for IgG, Fc gamma RI (CD
A:Reference number: A39878; MUID:91302383
A:Accession: A39878
A:Molecule type: DNA
A:Residues: 1-374 <VAN>

```

A:Cross-references: GB:M63830; GB:M63835; NID:g180277; PIDN:AAA35678.1; PID:g180279
 R:Porces, A.J.; Redecha, P.B.; Dobebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: I55577; MUID:93055454
 A:Accession: I70304
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:L03318; NID:g184840; PIDN:AAA36049.1; PID:g292169
 R:Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: A41357; MUID:89100284
 A:Accession: B41357
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', 26-337, 'T', 339-374 <ALL>
 A:Cross-references: GB:X14356; GB:M21091; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R:Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', 26-337, 'T', 339-374 <ALL>
 A:Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R:Perez, C.; Wietzerbin, J.; Benech, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A:Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
 sm.
 A:Reference number: I57525; MUID:93204964
 A:Accession: I57525
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RES>
 A:Cross-references: GB:S57204; NID:g298692; PIDN:AAD13887.1; PID:g4261587
 C:Genetics:
 A:Gene: GDB:FCGR1A; CD64
 A:Cross-references: GDB:135911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane F
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-29/Domain: extracellular #status predicted <EXT>
 F:117-170/Domain: immunoglobulin homology <IMM>
 F:293-313/Domain: transmembrane #status predicted <TM>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 61; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.00094;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EDGNTLRSPDL 12
 |||||
 Db 265 EDGNTLRSPDL 276
 |||||
 RESULT 3
 A46480
 Fc gamma (IgG) receptor high affinity mouse
 N:Alternate names: high affinity IgG receptor
 C:Species: Mus musculus (house mouse)
 C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
 C:Accession: A46480; A43511
 R:Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.
 J. Immunol. 148, 1570-1575, 1992
 A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI and
 A:Reference number: A46480; MUID:92166399
 A:Accession: A46480
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-404 <OSM>

A:Note: sequence extracted from NCBI backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212,
 R:Seares, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.
 J. Immunol. 144, 371-378, 1990
 A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for
 A:Reference number: A43511; MUID:90111035
 A:Accession: A43511
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-404 <SEA>
 A:Cross-references: GB:M31314; NID:g200752; PIDN:AAA40056.1; PID:g200753
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: immunoglobulin receptor; transmembrane protein
 F:127-179/Domain: immunoglobulin homology <IMM>

Query Match 82.0%; Score 50; DB 2; Length 404;
 Best Local Similarity 83.3%; Pred. No. 0.13;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 EDGNTLRSPDL 12
 |||||
 Db 274 EDSSVLRSPDL 285
 |||||

RESULT 4
 S55946
 Probable membrane protein YLR390W - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein L8084.10
 C:Species: Saccharomyces cerevisiae
 C:Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 06-Feb-1998
 C:Accession: S55946
 R:Du, Z.
 submitted to the EMBL Data Library, January 1995
 A:Description: The sequence of S. cerevisiae cosmid 8084.
 A:Reference number: S55944
 A:Accession: S55946
 A:Molecule type: DNA
 A:Residues: 1-112 <DUZ>
 A:Cross-references: EMBL:U19729; NID:6625097; PID:g625107; MIPS:YLR390W
 A:Experimental source: strain 5288C (AB972)
 C:Genetics:
 A:Gene: SGD:ECM19
 A:Cross-references: SGD:S0004382; MIPS:YLR390W
 A:Map position: 12R
 C:Keywords: transmembrane protein
 F:37-53/Domain: transmembrane #status predicted <TM>

Query Match 68.9%; Score 42; DB 2; Length 112;
 Best Local Similarity 66.7%; Pred. No. 1;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 EDGNTLRSPDL 12
 |||||
 Db 82 EDGNTLRSPDL 93
 |||||

RESULT 5
 S77922
 shaggy-like protein kinase Iota (EC 2.7.1.-) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Jul-1997 #sequence_revision 01-Aug-1997 #text_change 20-Jun-2000
 C:Accession: S77922
 R:Dornelas, M.C.; Lecharny, A.; Kreis, M.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: S77922
 A:Accession: S77922
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-407 <DOR>
 A:Cross-references: EMBL:X99696; NID:g1480077; PIDN:CAA68027.1; PID:g1480078
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP; phosphotransferase; protein kinase

F:68-329/Domain: protein kinase homology <KIN>
F:76-84/Region: protein kinase ATP-binding motif

Query Match 68.9%; Score 42; DB 2; Length 407;
Best Local Similarity 72.7%; Pred. No. 4.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DGNVLRKSPKL 12
||| ||| |||
Db 21 DGDALRKRPKL 31

RESULT 6

I48471
Fc gamma (19g) receptor high affinity - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: I48471
R:Prins, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarch, P.M.; Wicker, L.S.; Gaff
Science 260, 695-698, 1993
A:Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for Ig
A:Reference number: I48471; MUID:93242399
A:Accession: I48471
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-336 <RES>
A:Cross-references: EMBL:X70980; NID:g311748; PIDN:CAA50311.1; PID:g311749
C:Superfamily: Fc gamma receptor I; Immunoglobulin homology
C:Keywords: Immunoglobulin receptor
F:128-180/Domain: Immunoglobulin homology <IMM>

Query Match 67.2%; Score 41; DB 2; Length 336;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPKL 12
||| ||| |||
Db 279 EDSSVLRKSPKL 290

RESULT 7

H82339
conserved hypothetical protein VC0302 [Imported] - Vibrio cholerae (group O1 strain N169
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: H82339
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chadson, D.; Ermolaeva, M.D.; Yamathavan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R.
l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: H82339
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <HEI>
A:Cross-references: GB:AE004119; GB:AE003852; NID:g9654718; PIDN:AAF93476.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0302
A:Map position: 1
C:Superfamily: maltose permease

Query Match 65.6%; Score 40; DB 2; Length 396;
Best Local Similarity 63.6%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPKL 11
||| ||| |||
Db 383 ODGNVLRKSPKL 393

RESULT 8

D72605
probable long-chain-fatty-acid--CoA ligase (EC 6.2.1.3) APE1307 [similarity] - Aeropy
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 03-Nov-2000
C:Accession: D72605
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: D72605
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-651 <KAW>
A:Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAAR0298.1; PID:g5104984
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1307
C:Superfamily: Synchocystis long-chain-fatty-acid--CoA ligase; acetate--CoA ligase h
C:Keywords: acid-thiol ligase; coenzyme A
F:71-606/Domain: acetate--CoA ligase homology <ACL>

Query Match 65.6%; Score 40; DB 2; Length 651;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPKL 12
||| ||| |||
Db 432 EDGETLRSPAM 443

RESULT 9

T09481
mating type silencing regulator SIR4 - yeast (Kluyveromyces marxianus var. lactis)
N:Alternate names: silent information regulator 4
C:Species: Kluyveromyces marxianus var. lactis, Candida sphaerica
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T09481
R:Astrom, S.U.; Rine, J.
Genetics 148, 1021-1029, 1998
A:Title: Theme and variation among silencing proteins in Saccharomyces cerevisiae and
A:Reference number: Z16686; MUID:98198828
A:Accession: T09481
A:Status: translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1314 <AST>
A:Cross-references: EMBL:AF035007; NID:g3002787; PID:g3002789
A:Note: the source is designated as Kluyveromyces lactis
C:Genetics:
A:Gene: SIR4
C:Function:
A:Description: involved in both silencing and telomere length maintenance
C:Keywords: transcription regulation

Query Match 62.3%; Score 38; DB 2; Length 1314;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGNVLRKSPKL 11
||| ||| |||
Db 707 EDGNVSEKTPQ 717

RESULT 10

A25233
replication protein - Bacillus subtilis plasmid pIMJ3
C:Species: Bacillus subtilis
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 10-Dec-1999

Query Match	61.58;	Score 37.5;	DB 2;	Length 146;
Best Local Similarity	42.18;	Pred. No. 9.9;		
Matches	8;	Conservative	4;	Mismatches 0;
				Indels 7;
				Gaps 1.

RESULT 11

Query Match	61.5%	Score 37.5;	DB 2;	Length 158;
Best Local Similarity	42.1%;	Pred. No. 11;		
Matches	8; Conservative	4; Mismatches	0; Indels	7; Gaps 1;

RESULT 12

R; Dyke, K.G.H.; Curnock, S.P.
FEMS Microbiol. Lett. 58, 209-216, 1989

Query Match	61.5%	Score 37.5;	DB 2;	Length 158;
Best Local Similarity	42.1%;	Pred. No. 11;		
Matches	8;	Conservative	4;	Mismatches 0;
				Indels 7;
				Gaps 1

RESULT 13

A: Molecule type: DNA
A: Residues: 1-162 <LAMB>
A: Accession: B24497
A: Reference number: A91821; M01D:86304180
A: Genetics.

Query Match	61.5%;	Score 37.5;	DB 2;	Length 162;
Best Local Similarity	42.1%;	Pred. No. 11;		
Matches	8;	Conservative	4;	Mismatches 0;
				Indels 7;
				Gaps 1.

RESULT	14
S19534	

Query Match	60.7%	Score 37;	DB 2;	Length 181;
Best Local Similarity	63.6%;	Pred. No. 15;		
Matches	7;	Conservative	3;	Mismatches 1;
				Indels 0;
				Gaps 0;

RESULT 15
T05123

hypothetical protein F7H19.100 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: F05123
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; New
submitted to the Protein Sequence Database, July 1998
A:Reference number: Z15399
A:Accession: F05123
A:Molecule type: DNA
A:Residues: 1-268 <BEV>
A:Cross-references: EMBL:AL031018
A:Experimental source: cultivar Columbia; BAC clone F7H19
C:Genetics:
A:Map position: 4
A:Introns: 47/3; 105/3; 161/3
A>Note: F7H19.100

Query Match 60.7%; Score 37; DB 2; Length 268;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 2 DGNVLKRSPYL 12
|||:| |||
Db 170 DGNLNNYPYL 180

Search completed: June 4, 2001, 12:15:45
Job time: 278 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:07 ; Search time 37.56 Seconds
(without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-27

Sequence: 1 EDGNVLRKSPSL 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	374	1	FCG1_HUMAN
2	50	82.0	404	1	FCG1_MOUSE
3	42	68.9	407	1	KSG1_ARATH
4	37.5	61.5	146	1	REMA_BACSU
5	37.5	61.5	158	1	REMA_STRAU
6	37.5	61.5	162	1	REMA_STAEP
7	37	60.7	181	1	CRC1_HOMGA
8	37	60.7	388	1	SUCC_ECOLI
9	36	59.0	154	1	HS12_ORYSA
10	36	59.0	188	1	SSX3_HUMAN
11	36	59.0	216	1	ARA2_ARATH
12	36	59.0	428	1	GFAP_BOVIN
13	35	57.4	156	1	HS11_ARATH
14	35	57.4	156	1	HS12_ARATH
15	35	57.4	299	1	LAXC_LACTA
16	35	57.4	299	1	LAXP_LACTA
17	35	57.4	370	1	RECF_STRAU
18	35	57.4	376	1	OPPC_MYCPN
19	35	57.4	385	1	ACT2_PNECA
20	35	57.4	399	1	MTTE_HUMAN
21	35	57.4	434	1	IMPB_SALTY
22	35	57.4	437	1	PURA_VIBPA
23	35	57.4	662	1	YANK_RHISN
24	35	57.4	829	1	TOP1_XENLA
25	35	57.4	883	1	CHB_VIBHA
26	34.5	56.6	1256	1	MRP_STRSU
27	34	55.7	129	1	YBA2_ECOLI
28	34	55.7	161	1	HS11_CHERU
29	34	55.7	172	1	VP19_CRY
30	34	55.7	181	1	YE18_METJA
31	34	55.7	223	1	UBL1_HUMAN
32	34	55.7	223	1	UBL1_RAT
33	34	55.7	298	1	LAH1_SCHPO

34	34	55.7	565	1	KIR1_ECOLI	P08204	escherichia
35	34	55.7	569	1	KIR1_SALTY	P06188	salmonella
36	34	55.7	798	1	L100_ADEGX	P36856	avian adeno
37	34	55.7	922	1	YKFO_YEAST	P35736	saccharomyc
38	34	55.7	985	1	NAH1_YEAST	O99271	saccharomyc
39	34	55.7	1581	1	VGIP_BEV	P23052	berne virus
40	33	54.1	151	1	HS11_WHEAT	P12810	triticum ae
41	33	54.1	181	1	VP11_WHEAT	P28890	rice dwarf
42	33	54.1	189	1	VP11_RDVF	O85442	rice dwarf
43	33	54.1	215	1	SC4_OCTDO	P27012	octopus dof
44	33	54.1	252	1	YGRH_ECOLI	O46842	escherichia
45	33	54.1	291	1	YF92_SYNY3	P72970	synechocyst

ALIGNMENTS

```

RESULT 1
FCG1_HUMAN STANDARD: PRT: 374 AA.
ID FCG1_HUMAN
AC P12314; P12315;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA
DE R1) (FCRI) (IGG FC RECEPTOR I) (CD64).
GN FCGRIA OR FCGRI OR FCGI OR IGFRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89098339; PubMed=2974947;
RA Allen J.M., Seed B.;
RT "Nucleotide sequence of three cDNAs for the human high affinity FC
RL receptor (FCRI)".
RL Nucleic Acids Res. 16:11824-11824(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89100284; PubMed=2911749;
RA Allen J.M., Seed B.;
RT "Isolation and expression of functional high-affinity FC receptor
RL complementary DNAs."
RL Science 243:378-381(1989).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH
CC -1- AFFINITY RECEPTOR.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD64 entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd64.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X14356; CAA32537.1;
CC EMBL: X14355; CAA32536.1;
CC PIR: S03018; S03018.
CC PIR: S03019; S03019.
CC PIR: A41357; A41357.
CC PIR: B41357; B41357.
CC HSSP: P12319; IALT.
CC MIM: 146760;
CC InterPro: IPR003006;
CC Pfam: PF00047; 19; 3.

```

KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain; Alternative splicing; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 374
 FT DOMAIN 22 292
 FT TRANSMEM 293 313
 FT DOMAIN 314 374
 FT DOMAIN 36 92
 FT DOMAIN 117 175
 FT DOMAIN 205 267
 FT DISULFID 43 85
 FT DISULFID 124 168
 FT DISULFID 212 260
 FT CARBOHYD 59 59
 FT CARBOHYD 78 78
 FT CARBOHYD 152 152
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT VARSPLIC 333 374
 FT VARIANT 26 26
 FT VARIANT 59 59
 FT CONFLICT 25 25
 FT SEQUENCE 374 AA; 42605 MW; 2C2AA8103EC1656 CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 61; DB 1; Length 374;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EDGNVLRSPSL 12
 DB 265 EDGNVLRSPSL 276
 RESULT 2
 FCGL MOUSE STANDARD; PRT; 404 AA.
 AC P26151;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I).
 GN FCGR1 OR FCGL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90111035; PubMed=213686;
 RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RT "Molecular cloning and expression of the mouse high affinity Fc receptor for IgG."
 RT J. Immunol. 144:371-378(1990).
 RL [2]
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9216399; PubMed=1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RT "Structure and mapping of the gene encoding mouse high affinity Fc gamma RI and chromosomal location of the human Fc gamma RI gene."
 RL J. Immunol. 148:1570-1575(1992).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M31314; AAA40056.1;
 DR PIR: A43511; A43511.
 DR PIR: A4680; A4680.
 DR HSP: P12319; IALT.
 DR MGD: MGI:95498; Fcgr1.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; Ig; 3.
 KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 24
 FT CHAIN 25 404
 FT DOMAIN 25 297
 FT TRANSMEM 298 320
 FT DOMAIN 321 404
 FT DOMAIN 46 66
 FT DOMAIN 67 124
 FT DOMAIN 154 26
 FT CARBOHYD 28 28
 FT CARBOHYD 48 48
 FT CARBOHYD 69 69
 FT CARBOHYD 168 168
 FT CARBOHYD 249 249
 FT SEQUENCE 404 AA; 44887 MW; 1CAF0033842767E7 CRC64;
 Query Match
 Best Local Similarity 82.0%; Score 50; DB 1; Length 404;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 EDGNVLRSPSL 12
 DB 274 EDSSVLRSPSL 285
 RESULT 3
 KSGI ARATH STANDARD; PRT; 407 AA.
 AC Q39012;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SHAGGY-RELATED PROTEIN KINASE IOTA (EC 2.7.1.1-) (ASK-IOTA).
 GN ASK3 OR GSK1 OR T2D23.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 ON NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA; TISSUE=Root;
 RA Dornelas M.C., Schwebel-Dugue N., Thomas M., Lecharny A., Kreis M.;
 RT "Three new cDNAs related to SGG/GSK-3 (SHAGGY/glycogen synthase kinase-3) from Arabidopsis thaliana."
 RL (In) Plant Gene Register Pgr97-008.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Piao H.L., Jang H.J., Phn K.T., Lim J.H., Kang S.G., Jin J.B.,
 RA Hwang I.;
 RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]


```

RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA.
RA Liu S.X., Vaysberg M., Etyu P., Lee J.M., Lenz C., Pham P., Sakano H.,
RA Toriumi M., Yu G., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,
RA Altafi H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,
RA Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M.,
RA Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,
RA Federpiel N.A., Theologis A., to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY MEDIATE EXTRACELLULAR SIGNALS TO REGULATE
CC TRANSCRIPTION IN DIFFERENTIATING CELLS (BY SIMILARITY).
CC -1- PTM: AUTOPHOSPHORYLATED MAINLY ON THREONINE AND SERINE RESIDUES
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY; GSK-3 SUBSUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99696; CA68027.1;
DR EMBL: AF019927; AAB71545.1;
DR EMBL: AC068143; AAF82167.1;
DR HSPB: P24941; 1A01.
DR Mendel: 13957; AArch; 2319; 13957.
DR InterPro: IPR000719;
DR InterPro: IPR000290;
DR Pfam: PF00068; PKinase.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
KW Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation: Multigene family.
FT DOMAIN 70 354 PROTEIN KINASE.
FT NP_BIND 76 84 ATP (BY SIMILARITY).
FT BINDING 99 99 ATP (BY SIMILARITY).
FT ACT_SITE 195 195 BY SIMILARITY.
SQ SEQUENCE 407 AA; 46024 MW; 96BAC53754A764C CRC64;

OY 2 DGNVLRKSPRL 12.
DB 21 DGDALKRPEL 31.

RESULT 4
REMA_BACSU STANDARD; PRT; 146 AA.
AC PI3968;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE REPLICATION AND MAINTENANCE PROTEIN (PLASMID REPLICATION PROTEIN).
GN REPL.
OS Bacillus subtilis.
OC Plasmid pIM13.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86250583; PubMed=3087948;
RA Monod M., Denoya C., Dubnau D.;
RT "Sequence and properties of pIM13, a
RT macrocyclic lincosamide-streptogramin B resistance plasmid from

```

```

RT Bacillus subtilis".
RL J. Bacteriol. 167:138-147(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M13761; AAA98135.1;
DR PIR: A25233; A25233.
KW DNA replication; Plasmid.
SQ SEQUENCE 146 AA; 16617 MW; 9DF622ECBACA3BFF CRC64;

OY 1 EDGNVLRK-----SPRL 12
DB 101 EGNITKRRKTGVLMNPRL 119

RESULT 5
REMA_STADU STANDARD; PRT; 158 AA.
AC PI3969;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE REPLICATION AND MAINTENANCE PROTEIN (PLASMID REPLICATION PROTEIN);
GN REPL OR REP.
OS Staphylococcus aureus, and Staphylococcus simulans.
OC Plasmid pE5, and plasmid pPV142.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280; 1286;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=S. aureus; PLASMID=pE5;
RX MEDLINE=88032835; PubMed=2822666;
RA Projan S.V., Monod M., Narayanan C.S., Dubnau D.;
RT "Replication properties of pIM13, a naturally occurring plasmid found
RT in Bacillus subtilis, and of its close relative pE5, a plasmid native
RT to Staphylococcus aureus".
RL J. Bacteriol. 169:5131-5139(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. simulans; STRAIN=13044; PLASMID=pPV142;
RA Soukurt G.A., Solaiman D.K.Y., Steinberg D.H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17990; AAA98227.1;
DR EMBL: AF019140; AAC33147.1;
DR PIR: A29827; A29827.
KW DNA replication; Plasmid.
SQ SEQUENCE 158 AA; 18079 MW; 60E337A2A024455C CRC64;

Query Match 61.5%; Score 37.5; DB 1; Length 146;
Best Local Similarity 42.1%; Pred. No. 3.8;
Matches 8; Conservative 4; Mismatches 0; Indels 7; Gaps 1;

```

QY 1 EDGNVLR-----SPEL 12
 DB 101 EEGNIIKRTGVLMLNPEL 119

RESULT 6

REMA_STAEF STANDARD: PRT: 162 AA.

AC P06735: 052044:
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE REPLICATION AND MAINTENANCE PROTEIN (PLASMID REPLICATION PROTEIN).
 GN REPL.
 OS Staphylococcus epidermidis.
 OG plasmid pNE131.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=86304180; PubMed=3091582;
 RA Lampson B.C., Parisi J.T.;
 RT "Nucleotide sequence of the constitutive macrolide-lincosamide-
 streptogramin B resistance plasmid pNE131 from Staphylococcus
 RT epidermidis and homologues with staphylococcus aureus plasmids pEL194
 and pSN2".
 RT J. Bacteriol. 167:888-892(1986).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.

CC EMBL: M12730; AAA98295.1; -
 DR PIR: B24497; B24497.
 DR DNA replication; Plasmid.
 KW SEQUENCE 162 AA; 18447 MW; 09D4B8D5C5B957 CRC64;
 SQ

Query Match 61.5%; Score 37.5; DB 1; Length 162;
 Best Local Similarity 42.1%; Pred. No. 4.3;
 Matches 8; Conservative 4; Mismatches 0; Indels 7; Gaps 1;

QY 1 EDGNVLR-----SPEL 12
 DB 101 EEGNIIKRTGVLMLNPEL 119

RESULT 7

RCRL_HOMGA STANDARD: PRT: 181 AA.

AC P80029;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE CRUSTACEANIN C1 SUBUNIT.
 OS Homarus gammarus (European lobster) (Homarus vulgaris).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidae;
 OC Nephropidae; Nephropidae; Homarus.
 OX NCBI_TaxID=6707;
 RN [1]
 RP SEQUENCE.

RX MEDLINE=92037648; PubMed=1935978;
 RA Keen J.N., Caceres I., Eliopoulos E.E., Zagalsky P.F.,
 RA Findlay J.B.C.;
 RT "Complete sequence and model for the C1 subunit of the
 RT carotenoprotein, crustacyanin, and model for the dimer.

RT beta-crustacyanin, formed from the C1 and A2 subunits with
 RT astaxanthin."

RL Eur. J. Biochem. 202:31-40(1991).
 CC -1- FUNCTION: BINDS THE CAROTENOID ASTAXANTHIN WHICH PROVIDES THE BLUE
 CC COLORATION TO THE CARAPACE OF THE LOBSTER.
 CC -1- SUBUNIT: OLIGOMER; CAN FORM DIMERS (BETA-CRUSTACEANIN); OR
 CC COMPLEXES OF 16 SUBUNITS (ALPHA-CRUSTACEANIN). THERE ARE FIVE
 CC TYPES OF SUBUNITS: A1, A2, A3, C1 AND C2.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; CARAPACE.

CC -1- SIMILARITY: BELONGS TO THE LIPICALIN FAMILY.
 DR PIR: S19534; S19534.
 DR InterPro: IPR000566; -
 DR InterPro: IPR003057; -
 DR Pfam: PF00061; Lipocalin; 1.
 DR PRINTS: PR01273; INVERTICOLOR.
 DR PROSITE: PS00213; LIPOCALIN; FALSE-NEG.
 DR Pigment; Lipocalin; transport.
 FT DISULFID 12 121 BY SIMILARITY.
 FT DISULFID 51 173 BY SIMILARITY.
 SQ SEQUENCE 181 AA; 20667 MW; 6A05C7CBA3498BE8 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 181;
 Best Local Similarity 63.6%; Pred. No. 6;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 DGNVLRSPPEL 12
 DB 73 DGNVLRNGKL 83

RESULT 8
 SUCC_ECOLI STANDARD: PRT: 388 AA.

AC P07460;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SUCCINYL-CoA SYNTHETASE BETA CHAIN (EC 6.2.1.5) (SCS-BETA).

GN SUCC.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=86104124; PubMed=3002435;
 RA Buck D., Spencer M.E., Guest J.R.;
 RT "Primary structure of the succinyl-CoA synthetase of Escherichia
 RT coli."
 RT Biochemistry 24:6245-6252(1985).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RT Science 277:1453-1474(1997).

RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Samesi G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horinouchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map."

RL DNA Res. 3:137-155(1996).
 RP [4]
 RN SEQUENCE OF 1-12.
 RC STRAIN-K12/EMG2;
 RX MEDLINE-97443975; PubMed-9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 in the genome of *Escherichia coli* K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE-94193797; PubMed-8144675;
 RA Molodko W.T., Fraser M.E., James M.N.G., Bridger W.A.;
 RT "The crystal structure of succinyl-CoA synthetase from *Escherichia*
 coli at 2.5-A resolution.";
 RL J. Biol. Chem. 269:10883-10890(1994).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE-99141407; PubMed-9917402;
 RA Fraser M.E., James M.N., Bridger W.A., Molodko W.T.;
 RT "A detailed structural description of *Escherichia coli* succinyl-CoA
 synthetase.";
 RL J. Mol. Biol. 285:1633-1653(1999).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
 RX MEDLINE-20092606; PubMed-10625475;
 RA Joyce M.A., Fraser M.E., James M.N., Bridger W.A., Molodko W.T.;
 RT "ADP-binding site of *Escherichia coli* succinyl-CoA synthetase
 revealed by x-ray crystallography.";
 RL Biochemistry 39:17-25(2000).
 CC -1- CATALYTIC ACTIVITY: SUCCINATE + COA + ATP = SUCCINYL-COA + ADP +
 ORTHOPHOSPHATE.
 CC -1- ENZYME REGULATION: EXHIBITS TWO INTERESTING PROPERTIES: "SUBSTRATE
 SYNERGISM", IN WHICH THE ENZYME IS MOST ACTIVE FOR THE CATALYSIS
 OF ITS PARTIAL REACTIONS WHEN ALL THE SUBSTRATE BINDING SITES
 ARE OCCUPIED, AND "CATALYTIC COOPERATIVITY" BETWEEN ALTERNATING
 ACTIVE SITES IN THE TETRAMER, WHEREBY THE INTERACTION OF
 SUBSTRATES (PARTICULARLY ATP) AT ONE SITE IS NEEDED TO PROMOTE
 CATALYSIS AT THE OTHER.
 CC -1- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS.
 CC -1- MISCELLANEOUS: SUCCINYL-COA SYNTHETASE (SCS) OF *E. COLI* CATALYZES
 ITS REACTION VIA THREE STEPS THAT INVOLVE PHOSPHORYL ENZYME AND
 ENZYME-BOUND SUCCINYL PHOSPHATE AS INTERMEDIATES.
 CC -1- MISCELLANEOUS: DURING AEROBIC METABOLISM IT FUNCTIONS IN THE
 CITRIC ACID CYCLE, COUPLING THE HYDROLYSIS OF SUCCINYL-COA TO THE
 SYNTHESIS OF ATP & THUS REPRESENTS AN IMPORTANT SITE OF
 SUBSTRATE-LEVEL PHOSPHORYLATION. IT CAN ALSO FUNCTION IN THE
 OTHER DIRECTION FOR ANABOLIC PURPOSES, AND THIS MAY BE
 PARTICULARLY IMPORTANT FOR PROVIDING SUCCINYL-COA DURING ANAEROBIC
 GROWTH WHEN THE OXIDATIVE ROUTE FROM 2-OXOGLOUTARATE IS SEVERELY
 REPRESSED.
 CC -1- MISCELLANEOUS: THE BETA-SUBUNIT CONTAINS THE ATTACHMENT SITES FOR
 SUCCINATE. THE COMPLETE ACTIVE SITE IS PROBABLY LOCATED IN THE
 REGION OF ALPHA- BETA CONTACT.
 CC -1- SIMILARITY: TO OTHER BETA SUBUNITS OF SUCCINYL-COA SYNTHETASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J01619; AAA23899.1;
 DR EMBL; AE000176; AAC73822.1;
 DR EMBL; D90711; BAA35394.1;
 DR PIR; A24090; SYECSE.
 DR PDB; 1SCU; 20-APR-95.
 DR PDB; 2SCU; 02-AUG-99.
 DR PDB; 1CQJ; 10-JAN-00.
 DR PDB; 1CQJ; 10-JAN-00.
 DR SWISS-2DPAGE; P07460; COLL.

DR ECO2DBASE; E039.8; 6TH EDITION.
 DR Ecogene; EG10981; succ.
 DR InterPro; IPR000303;
 DR Pfam; PF00549; ligase-coa.1.
 DR PROSITE; PS01217; SUCCINYL-COA_LIG_3; 1.
 KW Ligase; Tricarboxylic acid cycle; 3d-structure.
 SQ SEQUENCE 388 AA; 41392 MW; 09C429EC97A823CF CRC64;
 Query Match 60.7%; Score 37; DB 1; Length 388;
 Best Local Similarity 63.6%; Pred. No. 14;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 DGNVLRSPSL 12
 DB 219 DGNVLRSPSL 229
 RESULT 9
 HS12_ORYSA STANDARD; PRT; 154 AA.
 ID HS12_ORYSA
 AC P31673;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE 17.4 KDA CLASS I HEAT SHOCK PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishi R., Hashimoto H., Uchimiya H., Kato A.;
 RL Submitted (JUL-1992) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBUNIT: FORMS OLIGOMERIC STRUCTURES.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 FAMILY.
 CC -1- SIMILARITY: PLANTS HAVE FOUR CLASSES OF HSP20: CLASS I AND II
 WHICH ARE CYTOPLASMIC; CLASS III WHICH IS IN THE CHLOROPLAST
 AND CLASS IV WHICH IS IN THE ENDOPLASMIC RETICULUM. THIS PROTEIN BELONGS
 TO CLASS I.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D12635; BAA02160.1;
 DR PIR; JS0710; JS0710.
 DR InterPro; IPR002068;
 DR Pfam; PF00011; HSP20.1.
 DR PROSITE; PS01031; HSP20.1.
 KW Heat shock; Multigene family.
 SQ SEQUENCE 154 AA; 17376 MW; B6A3C4C768598012 CRC64;
 Query Match 59.0%; Score 36; DB 1; Length 154;
 Best Local Similarity 72.7%; Pred. No. 7.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EDGNVLRSPSL 11
 DB 77 EDGNVLRSPSL 87
 RESULT 10
 SSX3_HUMAN STANDARD; PRT; 188 AA.
 ID SSX3_HUMAN

AC Q99909; 060223;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SSX3 PROTEIN.
 GN SSX3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis; and Fibrosarcoma;
 RX MEDLINE=96302330; PubMed=8697803;
 RA de Leeuw B., Balemans M., Geurts van Kessel A.;
 RT "A novel Krummel-associated box containing the SSX gene (SSX3) on the
 RT human X chromosome is not implicated in t(X;18)-positive synovial
 RT sarcomas".
 RL Cytogenet. Cell Genet. 73:179-183(1996).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98021352; PubMed=9378559;
 RA Gure A.O., Tureci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
 RA Knuth A., Pfeundschnub M., Old L.J., Chen Y.-T.;
 RT "SSX: a multigene family with several members transcribed in normal
 RT testis and human cancer".
 RL Int. J. Cancer 72:965-971(1997).
 CC -1- FUNCTION: COULD ACT AS A MODULATOR OF TRANSCRIPTION.
 CC -1- SIMILARITY: BELONGS TO THE SSX FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 KRAB-RELATED DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S82471; AAB37436.1; -;
 DR EMBL; 090840; AAC05819.1; -;
 DR PROSITE; PS50806; KRAB-RELATED; 1;
 KW Multigene family; Transcription regulation.
 FT DOMAIN 20 KRAB-RELATED.
 FT CONFLICT 95 83
 SO SEQUENCE 188 AA; 21682 MW; 368C7487C1450E72 CRC64;
 L -> Q (IN REF. 2).
 Query Match 59.0%; Score 36; DB 1; Length 188;
 Best Local Similarity 63.6%; Pred. No. 9.8;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EDGNYVKRSE 11
 DB 118 EGVNSKEVPE 128
 RESULT 11
 ID ARA2_ARATH STANDARD; PRT; 216 AA.
 AC P28185;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE RAS-RELATED PROTEIN ARA-2.
 GN ARA-2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=VARIOUS STRAINS; TISSUE=Leaf;
 RX MEDLINE=92084144; PubMed=1748311;
 RA Anal T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,
 RA Matsui M.;
 RT "Isolation and analysis of cDNAs encoding small GTP-binding proteins
 RT of Arabidopsis thaliana".
 RL Gene 108:259-264(1991).
 CC -1- SIMILARITY: TO RAS PROTEINS: BELONGS TO THE RAS SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D01024; BAA00829.1; -;
 DR PIR; J00907; J00907.
 DR PIR; J00639; J00639.
 DR HSSP; P28746; IAZK.
 DR InterPro; IPR001806;
 DR InterPro; IPR002380;
 DR Pfam; PF00071; ras; 1;
 DR PRINTS; PRO0449; RASTRNSPRMG.
 KW GTP-binding; Lipoprotein; Prenylation; Multigene family.
 FT NP_BIND 20 27 GTP (BY SIMILARITY).
 FT NP_BIND 68 72 GTP (BY SIMILARITY).
 FT NP_BIND 126 129 GTP (BY SIMILARITY).
 FT DOMAIN 42 50 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 213 213 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 214 214 GERANYL-GERANYL (BY SIMILARITY).
 SO SEQUENCE 216 AA; 23927 MW; 90744490A913954C CRC64;
 L -> Q (IN REF. 2).
 Query Match 59.0%; Score 36; DB 1; Length 216;
 Best Local Similarity 87.5%; Pred. No. 11;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDGNYVKR 8
 DB 203 EDGSVLRK 210
 RESULT 12
 ID GFAP_BOVIN STANDARD; PRT; 428 AA.
 AC Q28115;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE GLIAL FIBRILLARY ACIDIC PROTEIN, ASTROCYTE (GFAP).
 GN GFAP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Subcommissural organ;
 RA Bouchard P.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE OF 197-227 FROM N.A.
 RA Kirkpatrick B.W., Hart G.L., Moore S.S., Barendse W.;
 RT Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: GFAP, A CLASS-III INTERMEDIATE FILAMENT, IS A CELL-
 CC SPECIFIC MARKER THAT, DURING THE DEVELOPMENT OF THE CENTRAL
 CC NERVOUS SYSTEM, DISTINGUISHES ASTROCYTES FROM OTHER GLIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: Y08255; CA669422.1; -
 DR EMBL: L19867; AAA51413.1; -
 DR InterPro: IPR001664; -
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Heptad repeat pattern; Coiled coil.
 FT DOMAIN 1 68
 FT HEAD
 FT DOMAIN 69 373
 FT ROD
 FT DOMAIN 374 428
 FT TAIL
 FT DOMAIN 69 100
 FT COIL 1A.
 FT DOMAIN 101 111
 FT LINKER 1.
 FT DOMAIN 112 210
 FT COIL 1B.
 FT DOMAIN 211 226
 FT LINKER 12.
 FT DOMAIN 227 248
 FT COIL 2A.
 FT DOMAIN 249 252
 FT LINKER 2.
 FT DOMAIN 253 373
 FT COIL 2B.
 SQ SEQUENCE 428 AA; 49452 MM; 6891CBBBF2F59AD4 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 428;
 Best Local Similarity 60.0%; Pred. No. 24;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 DGNVLRKSP 11
 |||:|:|:
 Db 413 DGEVLRKSPQ 422

RESULT 13
 HS12_ARATH STANDARD; PRT; 156 AA.
 AC P19036;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE 17.4 KDA CLASS I HEAT SHOCK PROTEIN (HSP 17.4).
 GN HSP17.4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=90158498; PubMed=2482931;
 RA Takahashi T., Komeda Y.;
 RT "Characterization of two genes encoding small heat-shock proteins in
 Arabidopsis thaliana.";
 RL Mol. Gen. Genet. 219:365-372(1989).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 CC FAMILY.
 CC -1- SIMILARITY: PLANTS HAVE FOUR CLASSES OF HSP20: CLASS I AND II
 CC WHICH ARE CYTOPLASMIC; CLASS III WHICH IS IN THE CHLOROPLAST
 CC AND CLASS IV WHICH IS IN THE ENDOPLASMIC RETICULUM. THIS PROTEIN BELONGS
 CC TO CLASS I.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X17293; CAA5182.1; -
 DR PIR: J00351; J00351.
 DR InterPro: IPR002068; -
 DR Pfam: PF00011; HSP20; 1.
 DR PROSITE: PS01031; HSP20; 1.
 KW Heat shock; Multigene family.
 SQ SEQUENCE 156 AA; 17454 MM; F6FEA76EBA29501 CRC64;

Query Match 57.4%; Score 35; DB 1; Length 156;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRKSP 11
 |||:|:|:
 Db 79 EDGNVLRKSP 89

RESULT 14
 HS12_ARATH STANDARD; PRT; 156 AA.
 AC P13853;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE 17.6 KDA CLASS I HEAT SHOCK PROTEIN (HSP 17.6).
 GN HSP17.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA; TISSUE=Leaf;
 RX MEDLINE=90016890; PubMed=2798141;
 RA Helm K.W., Vierling E.;
 RT "An Arabidopsis thaliana cDNA clone encoding a low molecular weight
 heat shock protein.";
 RL Nucleic Acids Res. 17:7995-7995(1989).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 CC FAMILY.
 CC -1- SIMILARITY: PLANTS HAVE FOUR CLASSES OF HSP20: CLASS I AND II
 CC WHICH ARE CYTOPLASMIC; CLASS III WHICH IS IN THE CHLOROPLAST
 CC AND CLASS IV WHICH IS IN THE ENDOPLASMIC RETICULUM. THIS PROTEIN BELONGS
 CC TO CLASS I.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X16076; CAA34208.1; ALT_SEQ.
 DR PIR: S06074; S06074.
 DR InterPro: IPR002068; -
 DR Pfam: PF00011; HSP20; 1.
 DR PROSITE: PS01031; HSP20; 1.
 KW Heat shock; Multigene family.
 SQ SEQUENCE 156 AA; 17547 MM; 57BA4B7F1BD637DC CRC64;

Query Match 57.4%; Score 35; DB 1; Length 156;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRKSP 11
 |||:|:|:
 Db 80 EDGNVLRKSP 90

RESULT 15

LAXC LACIA

ID LAXC LACIA

STANDARD;

PRT; 299 AA.

AC PA2096;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, last sequence update)

DE LACX PROTEIN, CHROMOSOMAL.

GN LACX.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Lactococcus.

OX NCBI_TaxID=1360;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CNR2 270;

RX MEDLINE-92380489; PubMed-1339371;

RA Huang D.C.; Novel M.; Huang X.F.; Novel G.;

RT "Nonidentity between plasmid and chromosomal copies of ISS1-like

sequences in Lactococcus lactis subsp. lactis CNR270 and their

possible role in chromosomal integration of plasmid genes.";

RL Gene 118:39-46(1992).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

EMBL: X60456; CAA42987.1; 178BF4639FB1E2DB CRC64;

SEQUENCE 299 AA: 34491 MW; 178BF4639FB1E2DB CRC64;

Query Match 57.4%; Score 35; DB 1; Length 299;

Best Local Similarity 41.7%; Pred. No. 26;

Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDGNVLRSPET 12

Db 269 EGNTEDEKPCV 280

Search completed: June 4, 2001, 12:24:08
Job time: 570 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:23:28 ; Search time 107.68 Seconds

(without alignments)
13,062 Million cell updates/sec

Title: US-09-284-107-27

Perfect score: 61

Sequence: 1 EDGNVLRKSPSL 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP unclassified:*
13: SP vertebrate:*
14: SP virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	374	4	Q92663 homo sapien
2	61	100.0	375	4	Q92495
3	51	83.6	349	6	Q9M270
4	42	68.9	99	5	Q97172
5	42	68.9	112	3	Q06011
6	42	68.9	299	5	Q9NFA1
7	42	68.9	407	10	Q39012
8	40	65.6	396	2	Q9KVS5
9	40	65.6	651	1	Q9YCF0
10	38.5	63.1	542	2	Q9ZAS6
11	38	62.3	264	10	Q9SMK7
12	38	62.3	709	10	Q9MAX4
13	38	62.3	1314	3	Q59904
14	37.5	61.5	158	2	Q53734
15	37	60.7	268	10	Q82741
16	37	60.7	295	10	Q9SKD9
17	37	60.7	619	10	Q9XID4
18	37	60.7	767	10	Q9LOE8
19	37	60.7	782	10	Q9SSEA

20	37	60.7	871	5	Q17088	Q17088 caenorhabd1
21	37	60.7	1354	5	Q76251	Q76251 drosophila
22	37	60.7	1354	5	Q9V6K1	Q9V6K1 drosophila
23	37	60.7	1908	5	Q9V8K9	Q9V8K9 drosophila
24	37	60.7	1970	5	Q16103	Q16103 drosophila
25	36	59.0	88	10	Q9M412	Q9M412 quercus sub
26	36	59.0	105	10	Q9M414	Q9M414 quercus sub
27	36	59.0	110	10	Q9M419	Q9M419 quercus sub
28	36	59.0	110	10	Q9M417	Q9M417 quercus sub
29	36	59.0	110	10	Q9M416	Q9M416 quercus sub
30	36	59.0	110	10	Q9M411	Q9M411 quercus sub
31	36	59.0	110	10	Q9M410	Q9M410 quercus sub
32	36	59.0	110	10	Q9M409	Q9M409 quercus sub
33	36	59.0	110	10	Q9M408	Q9M408 quercus sub
34	36	59.0	110	10	Q9LDG9	Q9LDG9 q heat shoc
35	36	59.0	154	10	Q93439	Q93439 oryza sativ
36	36	59.0	154	10	Q9ZS24	Q9ZS24 castanea sa
37	36	59.0	154	10	Q9ZP84	Q9ZP84 quercus sub
38	36	59.0	159	10	Q40867	Q40867 pennisetum
39	36	59.0	160	10	P93440	P93440 oryza sativ
40	36	59.0	160	10	Q9SMF2	Q9SMF2 hellianthus
41	36	59.0	288	2	Q9XB56	Q9XB56 erwilia car
42	36	59.0	290	2	Q9PAH0	Q9PAH0 xylella fas
43	36	59.0	431	5	Q9NSD4	Q9NSD4 caenorhabd1
44	36	59.0	570	10	Q9LOE3	Q9LOE3 arabidopsis
45	36	59.0	648	5	Q9VXP2	Q9VXP2 drosophila

ALIGNMENTS

RESULT 1
Q92663 ID Q92663 PRELIMINARY; PRT; 374 AA.
AC Q92663;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE FC GAMMA RECEPTOR I.
GN A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93055454; PubMed=1430234;
RA Porjes A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RT "Novel Fc gamma receptor I family gene products in human mononuclear
cells.";
RT J. Clin. Invest. 90:2102-2109(1992).
DR EMBL: L03418; AAA36049.1; -;
DR HSSP: P12319; IALT.
DR INTERPRO: IPR003006; -;
DR PFM: PFM0047; 1g; 3.
DR PRODOM: PD002534; -; 1.
SQ SEQUENCE 374 AA; 42632 MW; D33D59398CEEA699 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGNVLRKSPSL 12
Db 265 EDGNVLRKSPSL 276

RESULT 2
ID Q92495 PRELIMINARY; PRT; 375 AA.

AC 092495;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR TYPE I (FC GAMMA RIB-FC GAMMA RECEPTOR)
 GN CD64 OR FC-GAMMA-RIB
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Benech P.D., Sastri K.N., Iyer R.R., Eichbaum Q.G., Raven D.,
 RA Ezekowitz A.B.,
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-224 FROM N.A.
 RA MEDLINE-93018627; PubMed=1402657;
 RA Benech P.D., Sastri K., Iyer R.R., Eichbaum Q.G., Raven D.P.,
 RA Ezekowitz R.A.,
 RT Definition of interferon gamma-response elements in a novel human FC
 RT gamma receptor gene (Fc gamma RIB) and characterization of the gene
 RT structure.
 RL J. Exp. Med. 176:1115-1123(1992)
 DR EMBL; M91555; AA58414.1;
 DR EMBL; M91550; AA58414.1; JOINED
 DR EMBL; M91551; AA58414.1; JOINED
 DR EMBL; M91552; AA58414.1; JOINED
 DR EMBL; M91553; AA58414.1; JOINED
 DR EMBL; M91554; AA58414.1; JOINED
 DR EMBL; S45709; AAD13842.1;
 DR EMBL; S45707; AAD13842.1; JOINED
 DR EMBL; S45708; AAD13842.1; JOINED
 DR EMBL; S45704; AAD13842.1; JOINED
 DR EMBL; S45705; AAD13842.1; JOINED
 DR HSSP; P12319; 1ALT
 DR INTERPRO: IPR003006;
 DR PIRAM; PF00047; 19; 3;
 DR PRODOM; PD002534; -; 1
 SQ SEQUENCE 375 AA; 42881 MW; A84D46AC70DD0F91 CRC64;
 Query Match 100.0%; Score 61; DB 4; Length 375;
 Best Local Similarity 100.0%; Pred. No. 0.0031;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EDGNVLRKSPSL 12
 DB 266 EDGNVLRKSPSL 277
 RESULT 3
 O9M2T0 PRELIMINARY; PRT; 349 AA.
 AC O9M2T0;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR I.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yan Y., Li X., Wang A., Zhang G.,
 RT Molecular cloning and identification of full-length cDNA encoding
 RT high affinity FC receptor for bovine IgG (Fc gamma RI).
 RL Vet. Immunol. Immunopathol. 75:151-159(2000).
 DR EMBL; AF162866; AAF80477.1;
 KW Receptor.

SQ SEQUENCE 349 AA; 39608 MW; D0B7B2EF9408C02 CRC64;
 Query Match 83.6%; Score 51; DB 6; Length 349;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 2 DGNVLRKSPSL 12
 DB 266 DGNVLRKSPSL 276
 RESULT 4
 O97172 PRELIMINARY; PRT; 99 AA.
 AC O97172;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE FCP3C PROTEIN.
 GN FCP3C OR EG:140G11.3.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OX Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE-20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burdits K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chertys J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT The genome sequence of Drosophila melanogaster.
 RL Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RA Borkova D., Minana B., Kafatos F.C.,
 RT "Sequencing the distal X chromosome of Drosophila melanogaster."
 RT Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RA Benos P.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003426; AAF45849.1;
 DR EMBL; AL035395; CAB37612.1;
 DR FLYBASE; FBgn0000644; Fcpc3C.
 SQ SEQUENCE 99 AA; 11150 MW; 0BF0050D1F11758 CRC64;

Query Match 68.9%; Score 42; DB 5; Length 99;
 Best Local Similarity 54.5%; Pred. No. 2.4;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EDGNVLRKSP 11
 Db 36 DDGNVIRKGP 46

RESULT 5
 ID 006011 PRELIMINARY; PRT; 112 AA.
 AC 006011;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE CHROMOSOME XII COSMID 8084.
 GN L8084.10.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C (AB972);
 RX MEDLINE=97313267; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansgorge W.,
 RA Benes V., Bruckner M., Delius H., Dubois E., Dusterhoft A.,
 RA Bentian K.D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Miesse H.W., Miosga T., Mostl D.,
 RA Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
 RA Muller-Auer S., Neutwich U., Obermaier B., Piravandi E., Pohl T.M.,
 RA Portetelle D., Purrelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urrestazu L.A., Vandenbol M., Verhasselt P.,
 RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
 RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hobeisel J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII."
 RL Nature 387:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C (AB972);
 RA Du Z.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C (AB972);
 RA Waterston R.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C (AB972);
 RA Cherry J.M.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U19729; AAB8349.1;
 SQ SEQUENCE 112 AA; 12515 MW; B469042AC2025AA7 CRC64;

Query Match 68.9%; Score 42; DB 3; Length 112;
 Best Local Similarity 56.7%; Pred. No. 2.7;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 EDGNVLRKSP 12

Db 82 EDGNVLRKSP 93

RESULT 6
 ID 09NF41 PRELIMINARY; PRT; 299 AA.
 AC 09NF41;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE BACN4124.C.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mottler, Cadieu, Dreano, Lelaure, Galibert F.;
 RT "Sequencing the distal X chromosome of Drosophila melanogaster."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Benos P.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL121805; CAB58072.1;
 SQ SEQUENCE 299 AA; 32007 MW; C7224AFCD0809B6 CRC64;

Query Match 68.9%; Score 42; DB 5; Length 299;
 Best Local Similarity 54.5%; Pred. No. 8;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EDGNVLRKSP 11
 Db 236 DDGNVIRKGP 246

RESULT 7
 ID 039012 PRELIMINARY; PRT; 407 AA.
 AC 039012;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE SHAGGY-LIKE PROTEIN KINASE IOTA (T2D23.9 PROTEIN).
 GN ASK IOTA OR GSK1 OR T2D23.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOTS; Schwebel-Dugue N., Thomas M., Leclary A., Kreis M.;
 RA Dornelas M.C.;
 RL Plant Physiol. 113:306-306(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Piao H.L., Jang H.J., Pih K.T., Lim J.H., Kang S.G., Jin J.B.,
 RA Hwang I.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Liu S.X., Vaysberg M., Etgu P., Lee J.M., Lenz C., Pham P., Sakano H.,
 RA Toriumi M., Yu G., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,
 RA Alcafi H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,
 RA Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M.,
 RA Nguyen M., Palm C.J., Shin P., Southwick A., Davis R.W., Ecker J.R.,
 RA Federpiet N.A., Theologis A.;

RT "The sequence of BAC T2D23 from Arabidopsis thaliana chromosome 1."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X99696; CA68027.1;
 DR EMBL: AF019927; AB071545.1;
 DR EMBL: AC068143; AAF82167.1;
 DR HSSP: P24941; JUSU.
 DR MENDEL: 13957; Arath:2319;13957.
 DR INTERPRO: IPR000719;
 DR INTERPRO: IPR002290;
 DR PFAM: PF00069; pkinase.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 SQ SEQUENCE 407 AA; 46024 MW; 96BC4C53754A764C CRC64;

Query Match 68.9%; Score 42; DB 10; Length 407;
 Best Local Similarity 72.7%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DGNVLRSPPEL 12
 DB 21 DGDALKRPEL 31

RESULT 8
 O9KV55 PRELIMINARY; PRT; 396 AA.
 AC O9KV55;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 GN HYPOTHETICAL PROTEIN VC0302.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE-20406833; PubMed-10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tellein H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Uitterlind T., Fleischmann R.D., Niernan W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.C., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae."
 RL Nature 406:477-483(2000).
 DR EMBL: AE004118; AAF93476.1;
 DR TIGR: VC0302;
 SQ SEQUENCE 396 AA; 43560 MW; 4D840465DBDC2757 CRC64;

Query Match 65.6%; Score 40; DB 2; Length 396;
 Best Local Similarity 63.6%; Pred. No. 25;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPPEL 11
 DB 383 QDGNALKAPE 393

RESULT 9
 O9YCF0 PRELIMINARY; PRT; 651 AA.
 AC O9YCF0;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE 651AA LONG HYPOTHETICAL LONG-CHAIN-FATTY-ACID--COA LIGASE.

GN APE1307.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
 CC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K1;
 RX MEDLINE-99310339; PubMed-10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
 RA Jinno K., Takahashi M., Sekine M., Baba S., Anai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamaya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K., Kubota K., Nakamura Y.,
 RA Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1."
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AP000061; BAA80298.1;
 DR HSSP: P08659; ILCT.
 DR INTERPRO: IPR000873;
 DR PROSITE: PS00455; AMP_BINDING; UNKNOWN_1.
 DR LIGASE.
 SQ SEQUENCE 651 AA; 74662 MW; D641C4B159CD6754 CRC64;

Query Match 65.6%; Score 40; DB 1; Length 651;
 Best Local Similarity 58.3%; Pred. No. 44;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 EDGNVLRSPPEL 12
 DB 432 EDGEILMRSPAM 443

RESULT 10
 O9ZAS6 PRELIMINARY; PRT; 542 AA.
 AC O9ZAS6;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 GN SUPRACE LIPOPROTEIN DPPA.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CS101;
 RX MEDLINE-98343814; PubMed-9680220;
 RA Podbielski A., Leonard B.A.;
 RT "The group A streptococcal dipeptide permease (Dpp) is involved in the
 RT uptake of essential amino acids and affects the expression of cysteine
 RT protease."
 RL Mol. Microbiol. 28:1323-1334(1998).
 DR EMBL: U78968; AAC67217.1;
 DR INTERPRO: IPR000914;
 DR PFAM: PF00496; SBP_dac_5; 1.
 KW Lipoprotein.
 SQ SEQUENCE 542 AA; 61145 MW; 9CB9460D84F71FB6 CRC64;

Query Match 63.1%; Score 38.5; DB 2; Length 542;
 Best Local Similarity 56.2%; Pred. No. 68;
 Matches 9; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

OY 2 DGNVLRSPPEL 12
 DB 66 ECNITHSLRSPPEL 81

```

RESULT 11
ID 09SMK7 PRELIMINARY; PRT; 264 AA.
AC 09SMK7;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE PUTATIVE ENOYL COA HYDRATASE.
OS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Cicer.
OX NCBI_TaxID=3827;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;
RA Labrador E., Esteban R., Dopico B.;
RT "An putative enoyl CoA hydratase expressed in Cicer arietinum
epicotyls."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: A075305; CAB61740.1; -
DR INTERPRO: IPR001753; -
DR PRAM: PF00378; ECH; 1.
SQ SEQUENCE 264 AA; 29066 MW; FF529E18AEB20739 CRC64;

Query Match
Best Local Similarity 62.3%; Score 38; DB 10; Length 264;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRSPEL 12
DB 185 EDGELKRSREI 196

RESULT 12
ID 09MAX4 PRELIMINARY; PRT; 709 AA.
AC 09MAX4;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE TRANSPOSASE.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-JITTERY;
RA Van X., Li J., Comperthwaite M., Fu H., Dooner H.;
RT "Jittery, a low-copy, Mu-related transposon apparently mobilized by
BSMV infection."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF247646; AAF6682.1; -
SQ SEQUENCE 709 AA; 81941 MW; F1080D806D982ABF CRC64;

Query Match
Best Local Similarity 62.3%; Score 38; DB 10; Length 709;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 EDGNVLRSPPE 11
DB 586 DDGNLLEKPE 596

RESULT 13
ID 059904 PRELIMINARY; PRT; 1314 AA.
AC 059904;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)

```

```

DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE SILENT INFORMATION REGULATOR 4.
CN STRA.
OS Kluyveromyces lactis.
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98198828; PubMed=9539421;
RA Astrom S.U., Rhine J.;
RT "Theme and variation among silencing proteins in Saccharomyces
cerevisiae and Kluyveromyces lactis."
RL Genetics 148:1021-1029(1998).
DR EMBL: AF035007; AAC39438.1; -
SQ SEQUENCE 1314 AA; 146734 MW; 3C78CBB6319F8A6 CRC64;

Query Match
Best Local Similarity 62.3%; Score 38; DB 3; Length 1314;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGNVLRSPPE 11
DB 707 EDGNVSEKTPQ 717

RESULT 14
ID 053734 PRELIMINARY; PRT; 158 AA.
AC 053734;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-MAY-1996 (TREMblrel. 10, Last annotation update)
DE PLASMID PSN2 INCLUDES AN UNKNOWN GENE.
OS Staphylococcus aureus.
OC Plasmid PSN2.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82119973; PubMed=7056699;
RA Khan S.A., Novick R.P.;
RT "Structural analysis of plasmid PSN2 in Staphylococcus aureus: no
involvement in enterotoxin B production."
RL J. Bacteriol. 149:642-649(1982).
DR EMBL: J01763; AAA91889.1; -
DR EMBL: V01282; CAA24595.1; -
KW Plasmid.
SQ SEQUENCE 158 AA; 18182 MW; EF0426F315C6F659 CRC64;

Query Match
Best Local Similarity 61.5%; Score 37.5; DB 2; Length 158;
Matches 8; Conservative 4; Mismatches 0; Indels 7; Gaps 1;

OY 1 EDGNVLR-----SPEL 12
DB 101 EGNVIRRRGALMLNP 119

RESULT 15
ID 082741 PRELIMINARY; PRT; 268 AA.
AC 082741;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE HYPOTHETICAL 30.1 KDA PROTEIN.
GN F7H19.100 OR ATAG22920.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

```

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Bancroft I.,
RA Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.,
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031018; CAl19807.1; -
DR EMBL; AL161558; CAB79247.1; -
KW Hypothetical protein.
SQ SEQUENCE 268 AA; 30052 MW; 1B3F128F18EA79B2 CRC64;

Query Match 60.7%; Score 37; DB 10; Length 268;
Best Local Similarity 63.6%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 DGNVLRKSPQL 12
|||:| |||
Db 170 DGNLNNYPQL 180

Search completed: June 4, 2001, 12:23:30
Job time: 593 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:32 ; Search time 106.14 Seconds
(without alignments)
6.463 Million cell updates/sec

Title: US-09-284-107-28
Sequence: 1 ELQVIGLQLPFP 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 390729 seqs, 57163235 residues
Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_0401.*
2: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	12	19	W60560
2	60	100.0	344	17	R91439
3	60	100.0	344	21	Y96183
4	60	100.0	374	17	R91438
5	60	100.0	374	17	W00859
6	60	100.0	374	19	W80448
7	60	100.0	374	19	W97833
8	60	100.0	374	19	W97834
9	60	100.0	374	21	Y96134
10	60	100.0	374	21	Y96226
11	60	100.0	399	21	B43683

12	52	86.7	377	20	W86195	Human FC receptor
13	48	80.0	344	13	R22549	Human macrophage-s
14	48	80.0	374	13	R20811	Human macrophage-s
15	48	80.0	374	13	R22550	Human macrophage-s
16	39	65.0	338	20	W88361	Human lymphocyte a
17	39	65.0	380	12	R13272	Polynhedin-soluble
18	39	65.0	422	20	W88360	Human lymphocyte a
19	39	65.0	498	12	R13270	Lymphocyte Activat
20	39	65.0	498	16	R87089	Human immunoglobul
21	38	63.3	164	16	R67587	Bovine male enhanc
22	38	63.3	174	16	R67586	Murine male enhanc
23	38	63.3	649	22	B65680	Novel protein kina
24	37	61.7	87	21	G26029	Zea mays protein f
25	37	61.7	100	21	G25954	Zea mays protein f
26	37	61.7	108	21	G41005	Zea mays protein f
27	37	61.7	124	21	G25953	Zea mays protein f
28	37	61.7	178	21	G36067	Zea mays protein f
29	37	61.7	264	21	G37436	Protein involved i
30	37	61.7	481	20	Y06293	Human transcriptio
31	36	60.0	73	19	W50259	Mouse BID truncate
32	36	60.0	195	19	W50256	Mouse BH3 interact
33	36	60.0	195	21	R84016	Amino acid sequenc
34	36	60.0	329	21	B12525	Human Maf protein
35	36	60.0	363	21	Y96748	A. terreus Lovc de
36	36	60.0	1098	21	B02009	Type V adenylyl cy
37	36	60.0	1184	14	R32882	Cardiac adenylyl c
38	36	60.0	1184	16	R78519	Cardiac adenylyl c
39	36	60.0	1261	20	Y34819	Human type V adeny
40	35	58.3	100	20	Y34819	Amino acid sequenc
41	35	58.3	347	15	R56530	Protein kinase (CK
42	35	58.3	347	16	R76626	Human HRR25-like c
43	35	58.3	377	15	R56529	Protein kinase (CK
44	35	58.3	377	16	R76625	Human HRR25-like c
45	35	58.3	526	9	P80986	Sequence of fusion

ALIGNMENTS

RESULT 1
W60560 standard; peptide: 12 AA.
W60560:
18-AUG-1998 (first entry)
Oligopeptide from extracellular domain of CD64.
Extracellular domain: CD64; identification; antibody;
Immunohistochemical; immunofluorescent analysis; detection;
cell transformation; mutation; anti; oncogene.
Synthetic.
W09815833-A1.
16-APR-1998.
07-OCT-1997; 97WO-NL00557.
08-OCT-1996; 96EP-0202791.
(UYUT-) RIJSGUNIV UTRCHT.
De Kruij CA, Logtenberg T;
WPI: 1998-240964/21.
Identifying peptide(s) binding specifically to protein target - by
expressing on phage surface and testing for binding to immobilised
oligo:peptide derived from the target, useful for, e.g. identifying
specific antibodies

XX Example 1; Page 29; 40pp; English.

CC Synthetic oligopeptides W60537-61 are derived from the extracellular
CC domain of CD64. They were synthesised on 25 polyethylene rods as
CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
CC affinity binding to the CD64-covered rods. The specification describes a
CC method for the identification of a peptide able to bind specifically to a
CC target protein. The method comprises displaying the peptide on the
CC surface of a replicable display package, synthesising oligopeptides
CC derived from the target protein on a solid phase, and testing for binding
CC between the peptide and oligopeptides. The method is used to screen large
CC peptide libraries, especially to detect antibodies, or their fragments,
CC that bind to cell markers or that can differentiate between different
CC forms of the same protein, including bispecific antibodies that bind to
CC two non-overlapping epitopes on the same monomeric antigen or two
CC epitopes on different molecules. The genes/oligonucleotides that encode
CC selected peptides can be isolated and used for recombinant production of
CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
CC immunofluorescent analysis, and also to detect cell transformation caused
CC by mutation in (anti)oncogenes.

XX Sequence 12 AA;

Query Match 100.0%; Score 60; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 12
| | | | | | | | | | | | | |
Db 1 elqyigqlptp 12

RESULT 2
R91439 ID R91439 standard; Protein; 344 AA.

AC R91439;

DT 30-OCT-1996 (first entry)

DE Human FCRI (CDNA clone p98 product).

KW Cell surface antigen; cloning; immunoselection; immunotherapy;
KW therapy; diagnosis; vector; FCRI; Fc receptor.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 58 /note= "amino acid 58 is leu in clone p135
FT translated product"

PN US506126-A.

XX 09-APR-1996.

XX 25-FEB-1988; 88US-0160416.

XX 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

XX 13-JUL-1989; 89US-0379076.

XX 13-JUL-1990; 90US-0553759.

XX 18-OCT-1993; 93US-0139273.

PA (GEHO) GEN HOSPITAL CORP.

XX Arufo A, Seed B;
XX WPI; 1996-200279/20.
XX N-PSDB; T14718.

PT Cloning of cDNA encoding cell surface antigen - useful for isolation
PT of diagnostic and therapeutic proteins

XX Example 10; Column 55-56; 79pp; English.

CC The amino acid sequence (R91439) of human FCRI was detd. from a
CC cDNA clone, p98 (T14718), obtd. from a cDNA library using an
CC immunoselection cloning method. FCRI is a high affinity receptor
CC for the Fc portion of IgG, normally located on cell surfaces of
CC macrophages. 2 other cDNA clones, p135 (T14717) and p90 (T14719),
CC coded for a variants of the FCRI sequence; the C-terminal sequence
CC of the p98 product is truncated compared with those of the p135
CC and p90 products (see also R91438 and W00859). A fusion protein
CC of FCRI and a receptor ligand will be helpful in increasing the
CC potency of antibodies in therapy.

XX Sequence 344 AA;

Query Match 100.0%; Score 60; DB 17; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 12
| | | | | | | | | | | | | |
Db 277 elqyigqlptp 288

RESULT 3

Y96183 ID Y96183 standard; Protein; 344 AA.

AC Y96183;

DT 19-DEC-2000 (first entry)

DE Human macrophage-specific FCRI.

KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
KW panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
KW immune disorder; infection; asthma; immune-complex disease;

KW amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 15 /note= "encoded by GCG"

FT Misc-difference 38 /note= "encoded by ACC"

FT Misc-difference 50 /note= "encoded by CCG"

FT Misc-difference 51 /note= "encoded by ACC"

FT Misc-difference 55 /note= "encoded by CAC"

FT Misc-difference 56 /note= "encoded by CAC"

FT Misc-difference 60 /note= "encoded by TCC"

FT Misc-difference 64 /note= "encoded by CCC"

FT Misc-difference 82 /note= "encoded by CAG"

FT Misc-difference 116 /note= "encoded by CAA"

FT Misc-difference 117 /note= "encoded by GAA"

FT Misc-difference 122 /note= "encoded by GCA"

FT Misc-difference 123 /note= "encoded by TTC"

FT Misc-difference 126 /note= "encoded by ACC"

```

CC protein of FCRI and a receptor ligand will be helpful to increase
CC the potencies of antibodies in therapy.
XX
XX Sequence      344 AA:
S0
Query Match                      100.0%; Score 60; DB 21; Length 344;
Best Local Similarity    100.0%; Prd. No. 0.0049;
Matches   12; Conservative     0; Mismatches    0; Indels    0; Gaps
QY          1 ELQVLGLOLEPT 12
           |||VVIIIIIII
Db         277 elqvlglqlptp 288

RESULT      4
R91438
ID          R91438 standard; Protein; 374 AA.
AC          R91438;
XX
XX          30-OCT-1996 (first entry)
XX
DE          Human FCRI (cDNA clone p135 product).
KW          Cell surface antigen; cloning; immunoselection; immunotherapy;
KM          therapy; diagnosis; vector; FCRI; Fc receptor.
XX
OS          Homo sapiens.
FH          Key
FT          Misc-difference 25
FT          /note= "amino acid 25 is Thr in clone p90 and p98
FT          translated products"
FT          Misc-difference 58
FT          /note= "amino acid 58 is Val in p90 clone
FT          translated product"
XX
XX          US5506126-A.
PN          PN
PD          PD
PD          09-APR-1996.
XX
PF          25-FEB-1988;      88US-0160416.
XX
XX          01-DEC-1992;      92US-0983647.
PR          25-FEB-1988;      88US-0160416.
PR          13-JUL-1989;      89US-0379076.
PR          13-JUL-1990;      90US-0553759.
PR          18-OCT-1993;      93US-0139273.
XX
PA          (GEHO ) GEN HOSPITAL CORP.
XX
PI          Aruffo A. Seed B.
XX
DR          WPI; 1996-200279/20.
DR          N-PDB; T14717.
XX
XX          Cloning of cDNA encoding cell surface antigen - useful for isolation
PT          of diagnostic and therapeutic proteins
XX
XX          Example 10; Column 55-56; 79pp; English.
XX
XX          The amino acid sequence (R91438) of human FCRI was detd. from a
CC          cDNA clone, p135 (T14717), obtd. from a cDNA library using an
CC          immunoselection cloning method. FCRI is a high affinity receptor
CC          for the Fc portion of IgG, normally located on cell surfaces of
CC          macrophages. Another isolated cDNA clone, p90 (T14719), coded
CC          for a variant (W00859) of the FCRI sequence, and a third clone, p98
CC          (T14718), coded for an FCRI (R91439) having a different C-terminal
CC          sequence. A fusion protein of FCRI and a receptor ligand will
CC          be helpful in increasing the potency of antibodies in therapy.
XX
X0          Sequence      374 AA:

```

Query Match 100.0%; Score 60; DB 17; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELQVIGLQLP 12
 Db 277 elqvlglqlptp 288

RESULT 5

W00859

ID W00859 standard; Protein; 374 AA.

AC W00859;

DT 30-OCT-1996 (first entry)

DE Human FCRI (cDNA clone p90 product).

KW Cell surface antigen; cloning; immunoselection; immunotherapy;

XX therapy; diagnosis; vector; FCRI; Fc receptor.

OS Homo sapiens.

PN US5506126-A.

PD 09-APR-1996.

PF 25-FEB-1988; 88US-0160416.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 18-OCT-1993; 93US-0139273.

PA (GENO) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

XX WPI; 1996-200279/20.

XX N-PSDB; T14719.

XX Cloning of cDNA encoding cell surface antigen - useful for isolation

XX of diagnostic and therapeutic proteins

XX Example 10; Column 55-56; 79pp; English.

XX The amino acid sequence (W00859) of human FCRI was detd. from a

XX cDNA clone, p90 (T14719), obtd. from a cDNA library using an

XX immunoselection cloning method. FCRI is a high affinity receptor

XX for the Fe portion of IgG, normally located on cell surfaces of

XX macrophages. Other cDNA clones (see also T14717-18) coded for

XX variants (R91438-39) of the FCRI sequence. A fusion protein of

XX FCRI and a receptor ligand will be helpful in increasing the

XX potency of antibodies in therapy.

XX Sequence 374 AA;

XX

XX

XX

XX

XX

ID W80448 standard; Protein; 374 AA.

AC W80448;

DT 07-JUN-1999 (first entry)

DE Human Fc receptor I.

KW Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;

XX cloning.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 2 /note= "encoded by TGG"

XX Misc-difference 23 /note= "encoded by CTC"

XX Misc-difference 44 /note= "encoded by GAC"

XX Misc-difference 45 /note= "encoded by CTG"

XX Misc-difference 60 /note= "encoded by CCC"

XX Misc-difference 77 /note= "encoded by AAT"

XX Misc-difference 85 /note= "encoded by TCC"

XX Misc-difference 99 /note= "encoded by CAA"

XX Misc-difference 103 /note= "encoded by CCC"

XX Misc-difference 141 /note= "encoded by GGC"

XX Misc-difference 159 /note= "encoded by AAC"

XX Misc-difference 171 /note= "encoded by ATG"

XX Misc-difference 176 /note= "encoded by GTC"

XX Misc-difference 256 /note= "encoded by GGC"

XX US5830731-A.

XX 03-NOV-1998.

XX 21-MAY-1997; 97US-0861205.

XX 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

XX 13-JUL-1989; 89US-0379076.

XX 13-JUL-1990; 90US-0553759.

XX 21-MAY-1997; 97US-0861205.

XX (GENO) GEN HOSPITAL CORP.

XX Aruffo A, Seed B;

XX WPI; 1998-609251/51.

XX N-PSDB; V63456.

XX New cloning vector and polylinker - based on existing sequences for

XX efficient cloning and expression of mammalian cDNA(s), especially

XX human lymphocyte antigenic sequences

XX Example 10; Column 53-54; 75pp; English.

XX This is the amino acid sequence of human Fc receptor I (FCRI), as

XX deduced from cDNA clone p135 (see V63456) isolated using a rapid

XX immunoselection cloning method from a cDNA library expressed in COS

XX cells. The cDNA library was constructed from polyA RNA of cells

XX from a single patient undergoing extracorporeal interleukin-2

CC induction therapy. Clones p90 (see X07372) and p98 (see X07373) were also obtained. A novel method for cloning cDNAs from mammalian expression libraries is based on transient expression of an antigen in eukaryotic cells and selection of cells expressing the antigen by adherence to an antibody-coated substrate. The method is useful for the isolation and cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell. It has been used to clone genes (see V63442-63) encoding cell surface antigens from mammalian lymphocytes (see W80440-55). The isolated genes can be expressed in a prokaryotic or eukaryotic host cells to produce the encoded protein. The invention also provides high efficiency expression vectors (see V63441 and V63444) which allow the generation of very large mammalian expression libraries. The purified genes and proteins are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders of animals, including humans.

CC Sequence 374 AA;

Query Match 100.0%; Score 60; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 12
| | | | | | | | | | | | | |
DB 277 ELQVIGLQLP 288

RESULT 7

ID W97833 standard; Protein: 374 AA.

AC W97833;

DT 07-JUN-1999 (first entry)

DE Human Fc receptor I.

FC receptor I; FCRI; cell surface antigen; lymphocyte; human; cloning.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 2 /note= "encoded by TGG"

FT Misc-difference 23 /note= "encoded by CTC"

FT Misc-difference 44 /note= "encoded by GAC"

FT Misc-difference 45 /note= "encoded by GAC"

FT Misc-difference 60 /note= "encoded by CMC"

FT Misc-difference 77 /note= "encoded by CCC"

FT Misc-difference 85 /note= "encoded by AAT"

FT Misc-difference 99 /note= "encoded by TCC"

FT Misc-difference 103 /note= "encoded by CAA"

FT Misc-difference 141 /note= "encoded by CCC"

FT Misc-difference 159 /note= "encoded by GGC"

FT Misc-difference 171 /note= "encoded by AAC"

FT Misc-difference 176 /note= "encoded by ATG"

FT Misc-difference 256 /note= "encoded by GNC"

FT /note= "encoded by GGC"

EN US5830731-A.

PD 03-NOV-1998.

PP 21-MAY-1997; 97US-0861205.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 21-MAY-1997; 97US-0861205.

PA (GENO) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

DR WPI: 1998-609251/51.

DR N-PSDB; X07372.

PS Example 10; Column 53-54; 75pp; English.

CC This is the amino acid sequence of human Fc receptor I (FCRI), as deduced from cDNA clone p90 (see X07372) isolated using a rapid immunoselection cloning method from a cDNA library expressed in COS cells. The cDNA library was constructed from poly(A) RNA of cells from a single patient undergoing extracorporeal interleukin-2 induction therapy. Clones p135 (see V63456) and p98 (see X07373) were also obtained. A novel method for cloning cDNAs from mammalian expression libraries is based on transient expression of an antigen in eukaryotic cells and selection of cells expressing the antigen by adherence to an antibody-coated substrate. The method is useful for the isolation and cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell. It has been used to clone genes (see V63442-63) encoding cell surface antigens from mammalian lymphocytes (see W80440-55). The isolated genes can be expressed in a prokaryotic or eukaryotic host cells to produce the encoded protein. The invention also provides high efficiency expression vectors (see V63441 and V63444) which allow the generation of very large mammalian expression libraries. The purified genes and proteins are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders of animals, including humans.

CC Sequence 374 AA;

Query Match 100.0%; Score 60; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 12
| | | | | | | | | | | | | |
DB 277 ELQVIGLQLP 288

RESULT 8

ID W97834 standard; Protein: 374 AA.

AC W97834;

DT 07-JUN-1999 (first entry)

DE Human Fc receptor I.

FC receptor I; FCRI; cell surface antigen; lymphocyte; human;

```

KW cloning.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "encoded by TGC"
FT Misc-difference 23
FT /note= "encoded by CTC"
FT Misc-difference 44
FT /note= "encoded by GAC"
FT Misc-difference 45
FT /note= "encoded by CTG"
FT Misc-difference 60
FT /note= "encoded by CCC"
FT Misc-difference 77
FT /note= "encoded by AAT"
FT Misc-difference 85
FT /note= "encoded by TCC"
FT Misc-difference 99
FT /note= "encoded by CAA"
FT Misc-difference 103
FT /note= "encoded by CCC"
FT Misc-difference 141
FT /note= "encoded by GGC"
FT Misc-difference 159
FT /note= "encoded by AAC"
FT Misc-difference 171
FT /note= "encoded by ATG"
FT Misc-difference 176
FT /note= "encoded by GTC"
FT Misc-difference 256
FT /note= "encoded by GCG"
FT
XX
XX US5830731-A.
XX
XX 03-NOV-1998.
XX
XX
XX 21-MAY-1997; 97US-0861205.
XX
XX
XX 01-DEC-1992; 92US-0983647.
XX
XX 25-FEB-1988; 88US-0160416.
XX
XX 13-JUL-1989; 89US-0379076.
XX
XX 13-JUL-1990; 90US-0553759.
XX
XX 21-MAY-1997; 97US-0861205.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX
XX Aruffo A, Seed B;
XX
XX WPI: 1998-609251/51.
XX
XX N-PSDB: X07373.
XX
XX
XX New cloning vector and polylinker - based on existing sequences for
XX efficient cloning and expression of mammalian cDNA(s), especially
XX human lymphocyte antigenic sequences
XX
XX
XX Example 10; Column 53-54; 75pp; English.
XX
XX This is the amino acid sequence of human Fc receptor I (FcRI), as
XX deduced from cDNA clone P98/X2 (see X07373) isolated using a rapid
XX immunoselection cloning method from a cDNA library expressed in COS
XX cells. The cDNA library was constructed from polyA RNA of cells
XX from a single patient undergoing extracorporeal Interleukin-2
XX induction therapy. Clones p135 (see V63456) and p90 (see X07372)
XX were also obtained. A novel method for cloning cDNAs from mammalian
XX expression libraries is based on transient expression of an antigen
XX in eukaryotic cells and selection of cells expressing the antigen by
XX adhesion to an antibody-coated substrate. The method is useful for
XX the isolation and cloning of any protein which can be expressed and
XX transported to the cell surface membrane of a eukaryotic cell. It
XX has been used to clone genes (see V63442-63) encoding cell surface
XX antigens from mammalian lymphocytes (see W80440-55). The isolated

```

```

CC genes can be expressed in a prokaryotic or eukaryotic host cells to
CC produce the encoded protein. The invention also provides high
CC efficiency expression vectors (see V63441 and V63444) which allow
CC the generation of very large mammalian expression libraries. The
CC purified genes and proteins are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders of
CC animals, including humans.
XX
XX Sequence 374 AA:
SQ
Query Match 100.0%; Score 60; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ELOVIGLOLPTP 12
Db 277 ELQVIGLQPTP 288
RESULT 9
ID Y96134 standard; Protein; 374 AA.
AC Y96134;
XX
XX 19-DEC-2000. (first entry)
XX
XX Human macrophage-specific FcRI.
XX
XX Macrophage; FcRI; cell surface antigen; human; immunoselection;
XX panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;
XX immune disorder; infection; asthma; immune-complex disease;
XX amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Misc-difference 15 /note= "encoded by GCG"
XX Misc-difference 38 /note= "encoded by ACC"
XX Misc-difference 50 /note= "encoded by CGG"
XX Misc-difference 51 /note= "encoded by ACC"
XX Misc-difference 55 /note= "encoded by CAC"
XX Misc-difference 56 /note= "encoded by TCC"
XX Misc-difference 60 /note= "encoded by CCC"
XX Misc-difference 64 /note= "encoded by CAG"
XX Misc-difference 82 /note= "encoded by CAA"
XX Misc-difference 116 /note= "encoded by GAA"
XX Misc-difference 117 /note= "encoded by GCA"
XX Misc-difference 122 /note= "encoded by TGC"
XX Misc-difference 123 /note= "encoded by ACC"
XX Misc-difference 126 /note= "encoded by CCG"
XX Misc-difference 129 /note= "encoded by CAT"
XX Misc-difference 134 /note= "encoded by AAT"
XX Misc-difference 136 /note= "encoded by GTT"

```

FT Misc-difference 139. /note- "encoded by CCA"
 FT Misc-difference 140. /note- "encoded by AAT"
 FT Misc-difference 213. /note- "encoded by CAA"
 FT Misc-difference 216. /note- "encoded by TTC"
 FT Misc-difference 220. /note- "encoded by CGT"
 FT Misc-difference 268. /note- "encoded by AAT"
 FT Misc-difference 305. /note- "encoded by GTG"
 FT Misc-difference 306. /note- "encoded by AAC"
 FT Misc-difference 332. /note- "encoded by GGT"
 FT US6111093-A.
 PD 29-AUG-2000.
 PE 28-OCT-1998; 98US-0181612.
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 23-MAR-1990; 90US-0498809.
 PR 13-JUL-1990; 90US-0553759.
 XX (GEHO) GEN HOSPITAL CORP.
 XX Stamenkovic I, Seed B;
 XX WPI: 2000-586382/55.
 DR N-PSDB: A50592.
 PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 PT useful for immunodiagnosis and immunotherapy of immune-mediated
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 PT diseases
 XX Example 10; Column 53-55; 75pp; English.
 PS The present sequence is that of a human macrophage specific FCRI,
 CC as deduced from cDNA clone p135 (see A50592), which was isolated
 CC from a cDNA library expressed in COS cells using a novel method of
 CC the invention designed to isolate CSA nucleic acids. The method is
 CC based upon transient expression of a CSA in eukaryotic cells and
 CC physical selection of cells expressing the antigen by adhesion to
 CC (panning on) an antibody-coated substrate such as a culture dish.
 CC CSA nucleic acids isolated by the method of the invention, and the
 CC proteins they encode, are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders in
 CC animals, including humans. These disorders include asthma,
 CC immune-complex disease, amyloidosis, parasitic diseases or multiple
 CC sclerosis. FCRI is a high affinity receptor for the Fc portion of
 CC IgG, normally located on the cell surfaces of macrophages. The
 CC ability to interfere with such bonding, or to cause it to occur on
 CC surfaces other than macrophages, is useful in therapy. A fusion
 CC protein of FCRI and a receptor ligand will be helpful to increase
 CC the potencies of antibodies in therapy.
 CC
 XX Sequence 374 AA:

Query Match 100.0%; Score 60; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ELQVIGIQLPPT 12

Db 277 elqyvgiqp 288
 RESULT 10
 ID Y96226
 Y96226 standard; Protein; 374 AA.
 AC Y96226;
 DT 11-SEP-2000 (first entry)
 DE Human high affinity Fc receptor, FcgammaRI.
 KW Human; high affinity Fc receptor; FcgammaRI; immunoglobulin;
 KW infection; immune response; CD64; monocyte; macrophage; neutrophil;
 KW eosinophil; HIV; IgG; immunosuppressive; antineoplastic; cytostatic;
 KW antiinflammatory; AIDS; SLE; multiple myeloma; rheumatoid arthritis;
 KW systemic lupus erythematosus; tumour.
 XX Homo sapiens.
 OS
 XX EPI006183-A1.
 PN 07-JUN-2000.
 PD 03-DEC-1998; 98EP-0122969.
 PE 03-DEC-1998; 98EP-0122969.
 PR 03-DEC-1998; 98EP-0122969.
 XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX WPI: 2000-367968/32.
 DR N-PSDB: A27466.
 PT Novel Fc receptor lacking transmembrane domains, a signal peptide, and
 PT glycosylation, useful for diagnosing and treating immune disorders and
 PT cancer
 XX
 XX Disclosure; Page 26-28; 60pp; English.
 PS The present sequence is the human high affinity Fc receptor, FcgammaRI.
 CC FcgammaRI is also known as CD64. Fc receptors play an important
 CC role in defending the body against infections. First, pathogens are
 CC opsonised by serum immunoglobulins. The resulting complex then binds to
 CC cells expressing Fc receptors. FcgammaRI molecules are expressed by
 CC monocytes and macrophages, but expression can also be induced on
 CC neutrophils and eosinophils. Upon Fc receptor activation, immune effector
 CC pathways are activated, leading to immune response. The present sequence
 CC may be modified to produce recombinant versions. The recombinant Fc
 CC receptor consist only of the extracellular portion of the receptor and
 CC are not glycosylated i.e. they do not have transmembrane domains or
 CC signal peptides. The recombinant proteins may be used in immunoassays to
 CC determine the immune status of patients with chronic diseases of the
 CC immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple
 CC myeloma (MM), or rheumatoid arthritis. In addition, pharmaceutical
 CC compositions containing recombinant proteins may be used to treat or
 CC prevent autoimmune diseases, allergies or tumours, especially AIDS,
 CC rheumatoid arthritis or MM.
 XX Sequence 374 AA:

Query Match 100.0%; Score 60; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ELQVIGIQLPPT 12
 Db 277 elqyvgiqp 288
 RESULT 11

B43683
ID B43683 standard; Protein: 399 AA
AC B43683;
XX
XX
DT 08-FEB-2001 (first entry)
XX
XX

Human cancer associated protein sequence SEQ ID NO:1128.

Human; cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytostatic; proliferative; vulnery; immunomodulator; antidiabetic; antitumor; antineoplastic; antihypertensive; antiviral; antiinflammatory; antiallergic; antibacterial; coagulant; dermatologic; vasotrophic; antiproliferative; thrombolytic; coagulant; neotropic; immune disorder; haematopoietic cell disorder; autoimmune disorder; allergic reaction; graft versus host disease; organ rejection; haemostatic; thrombolytic; cardiovascular disorder; infection; neurological disease; drug screening.

OS Homo sapiens.

XX WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI: 2000-587533/55.

XX N-PSDB; C77892.

XX Novel isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer -

XX Claim 11: Page 1739-1740; 2352pp; English.

XX C77607 to C78448 encode the human cancer associated proteins given in
CC B43398 to B44239. The proteins can have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC cytostatic; proliferative; vulnery; immunomodulator; antidiabetic;
CC antitumor; antineoplastic; antihypertensive; antiviral; dermatologic;
CC antiallergic; antibacterial; coagulant; neotropic; immune disorder;
CC vasotrophic; antiproliferative; thrombolytic; coagulant; autoimmune disorder;
CC polypeptides can be used for preventing, treating or ameliorating medical
CC conditions and diagnosing pathological conditions. Polynucleotides and
CC polypeptides, antibodies, agonists and antagonists from the present
CC invention may be used to treat immune disorders by activating or
CC inhibiting the proliferation, differentiation or mobilisation of immune
CC cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC allergic reactions, graft versus host disease and organ rejection,
CC modulate haemostatic or thrombolytic activity, modulate inflammation,
CC cancers, cardiovascular disorders, neurological disease and bacterial or
CC viral infections. The peptides, nucleotides, antibodies, agonists and
CC antagonists may be also be used in drug screens. C78449 to C78457 and
CC B44240 represent sequences used in the exemplification of the present
CC invention.
XX

XX Sequence 399 AA;

XX Query Match 100.0%; Score 60; DB 21; Length 399;
XX Best Local Similarity 100.0%; Pred. No. 0.0057;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 ELQVGLQPLTP 12
XXXXXXXXXXXX

Do 302 elqyglqpltp 313

RESULT 12

W86195

ID W86195 standard; Protein: 377 AA.

XX W86195;

XX 10-MAY-1999 (first entry)

XX Human Fc receptor I.

XX Fc receptor I; CD36; cell surface antigen; human; cDNA library.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Protein 1..377

XX /note- "this sequence contains a considerable
XX number of differences from the sequence
XX deduced from the DNA sequence given in
XX the specification."

XX US5849898-A.

XX 15-DEC-1998.

XX 07-JUN-1995; 95US-0485447.

XX 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

XX 13-JUL-1989; 89US-0379076.

XX 23-MAR-1990; 90US-0498809.

XX 13-JUL-1990; 90US-0553759.

XX 07-JUN-1995; 95US-0485447.

XX (GENE) GEN HOSPITAL CORP.

XX Allen J, Amiot M, Aruffo A, Camerini D, Lauffer L;

XX Oquendo C, Seed B, Simmons D, Stamenkovic I, Stengelin S;

XX WPI: 1999-069813/06.

XX N-PSDB; V81213.

XX cDNA encoding human CD40 antigen - useful for cloning cDNA encoding

XX cell surface antigens, constructing cDNA libraries, expression

XX vectors for expression in eukaryotic cells or their fragments

XX Example 10; Column 55-56; 79pp; English.

XX This polypeptide human Fc receptor I (FcRI). FcRI cDNAs (see
XX V81213) were isolated using a rapid immunoselection cloning
XX method from a cDNA library expressed in COS cells. The cDNA
XX library was constructed from polyA RNA of cells from a single
XX patient undergoing extracorporeal interleukin-2 induction
XX therapy. DNA sequence analysis revealed that the cDNAs encoded
XX type I integral membrane proteins with 3 extracellular
XX immunoglobulin domains. The invention provides a novel method for
XX cloning cDNAs from mammalian expression libraries. This is based on
XX transient expression of an antigen in eukaryotic cells and physical
XX selection of cells expressing the antigen by adhesion to an
XX antibody-coated substrate. The method is useful for the isolation
XX and molecular cloning of any protein which can be expressed and
XX transported to the cell surface membrane of a eukaryotic cell.
XX CD40 cDNA (see V81198) is specifically claimed.

XX Sequence 377 AA;

XX Query Match 86.7%; Score 52; DB 20; Length 377;
XX Best Local Similarity 91.7%; Pred. No. 0.15;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 12
 |||||
 DB 280 elqvlgqlp 291

RESULT 13

R22549
 ID R22549 standard; Protein; 344 AA.

AC R22549;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FCRI receptor encoded by clone p98.

KW Rapid immunoselection cloning technique; cell surface antigen;
 immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN W09201049-A.

PD 23-JAN-1992.

PE 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PS (GEHO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amiot M;

DR WPI; 1992-056864/07.

DR N-PSDB; Q21179.

PT New CD53 cell surface antigen and DNA encoding it - for
 immuno-therapy and diagnosis of hematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p98/X2. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC while clone p98/X2 predicts a Thr residue. At position 58, p90
 CC (see Q21180) predicts Val and p98/X2 predicts Leu. Sequences
 CC predicted from all 3 clones show the typical features of a type I
 CC integral membrane protein and include a short hydrophobic signal
 CC sequence, a single 21-residue hydrophobic membrane-spanning domain,
 CC and a short, highly charged cytoplasmic domain. The extracellular
 CC portion contains six potential N-linked glycosylation sites and six
 CC Cys residues distributed among three C2 set Ig-related domains.
 CC A fusion protein of FCRI and a receptor ligand will be helpful to
 CC increase the potency of antibodies in therapy.

SO Sequence 344 AA;

Query Match 80.0%; Score 48; DB 13; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 10
 |||||
 DB 277 elqvlgqlp 286

RESULT 14

R20811
 ID R20811 standard; Protein; 374 AA.

AC R20811;

XX

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FCRI receptor encoded by clone p135.

KW Rapid immunoselection cloning technique; cell surface antigen;
 immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN W09201049-A.

PD 23-JAN-1992.

PE 15-JUL-1990; 90WO-US04986.

PR 13-JUL-1990; 90US-0553759.

PS (GEHO-) GEN HOSPITAL CORP.

PI Seed B, Aruffo A, Amiot M;

DR WPI; 1992-056864/07.

DR N-PSDB; Q21178.

PT New CD53 cell surface antigen and DNA encoding it - for
 immuno-therapy and diagnosis of hematopoietic neoplasms, etc.

PS Example 10; Page 94a; 160pp; English.

CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p135. It differs from the sequence predicted from
 CC clones p98/X2 and p90 (see Q21179 and Q21180, respectively) at
 CC position 25; p135 encodes a Ser residue and the other two clones
 CC predict a Thr residue. At position 58, p135 predicts Leu and p90
 CC predicts Val. Sequences predicted from all 3 clones show the
 CC typical features of a type I integral membrane protein and include
 CC a short hydrophobic signal sequence, a single 21-residue
 CC hydrophobic membrane-spanning domain, and a short, highly charged
 CC cytoplasmic domain. The extracellular portion contains six
 CC potential N-linked glycosylation sites and six Cys residues
 CC distributed among three C2 set Ig-related domains. A fusion protein
 CC of FCRI and a receptor ligand will be helpful to increase the
 CC potency of antibodies in therapy.

SO Sequence 374 AA;

Query Match 80.0%; Score 48; DB 13; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.77;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 10
 |||||
 DB 277 elqvlgqlp 286

RESULT 15

R22550
 ID R22550 standard; Protein; 374 AA.

AC R22550;

DT 21-MAY-1992 (first entry)

DE Human macrophage-specific FCRI receptor encoded by clone p90.

KW Rapid immunoselection cloning technique; cell surface antigen;
 immunodiagnosis; high affinity receptor.

OS Homo sapiens.

PN W09201049-A.

XX

PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90MO-US04986.
 XX
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO-) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Amiot M;
 XX
 DR WPI: 1992-056864/07.
 XX
 DR N-PSDB; Q21180.

PT New CD53 cell surface antigen and DNA encoding it - for
 XX immunotherapy and diagnosis of haematopoietic neoplasms, etc.
 PS
 XX

Example 10: Page 94a; 160pp; English.

CC This amino acid sequence was predicted from the cDNA sequence of
 CC cDNA clone p90. It differs from the sequence predicted from
 CC clone p135 (see Q21178) at position 25; p135 encodes a Ser residue
 CC and clone p90 predicts a Thr residue. At position 58, p90 predicts
 CC Val while clones p135 and p98/X2 (see Q21179) both predict Leu.
 CC Sequences predicted from all 3 clones show the typical features of
 CC a type I integral membrane protein and include a short hydrophobic
 CC signal sequence, a single 21-residue hydrophobic membrane-spanning
 CC domain, and a short, highly charged cytoplasmic domain. The
 CC extracellular portion contains six potential N-linked glycosylation
 CC sites and six Cys residues distributed among three C2 set Ig-related
 CC domains. A fusion protein of pC21 and a receptor ligand will be
 CC helpful to increase the potency of antibodies in therapy.
 XX

SQ Sequence 374 AA:

Query Match 80.0%; Score 48; DB 13; Length 374;
 Best Local Similarity 100.0%; Pred. NO. 0.77;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLP 10
 |||||
 Db 277 elqyIqlqIp 286

Search completed: June 4, 2001, 12:13:33
 Job time: 206 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:32 ; Search time 58.74 Seconds
(without alignments)

3.925 Million cell updates/sec

Title: US-09-284-107-28

Perfect score: 60

Sequence: 1 ELQVGLQLPPT 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	65.0	380	1	US-08-416-478A-6
2	39	65.0	380	2	US-08-474-988B-6
3	39	65.0	380	2	US-08-394-442B-6
4	39	65.0	470	1	US-08-416-478A-7
5	39	65.0	470	2	US-08-474-988B-7
6	39	65.0	470	2	US-08-394-442B-7
7	39	65.0	471	1	US-08-416-478A-2
8	39	65.0	471	2	US-08-474-988B-2
9	39	65.0	471	2	US-08-394-442B-2
10	39	65.0	476	2	US-08-737-271-1
11	39	65.0	476	4	US-09-058-555-1
12	39	65.0	498	1	US-08-416-478A-9
13	39	65.0	498	2	US-08-474-988B-9
14	39	65.0	498	2	US-08-394-442B-9
15	39	65.0	73	2	US-08-706-741B-50
16	39	65.0	73	2	US-08-924-695A-50
17	39	65.0	195	2	US-08-706-741B-6
18	39	65.0	195	2	US-08-924-695A-6
19	39	65.0	1098	1	US-08-726-214-10
20	39	65.0	347	1	US-08-454-097-33
21	39	65.0	347	3	US-08-185-359-33
22	39	65.0	361	1	US-08-415-751-36
23	39	65.0	377	1	US-08-454-097-31
24	39	65.0	377	1	US-08-185-359-31
25	39	65.0	256	2	US-08-030-096-2
26	39	65.0	256	2	US-08-481-658B-51
27	39	65.0	256	2	US-08-477-504A-51

28	33	55.0	256	2	US-08-486-756A-51	Sequence 51, Appl
29	33	55.0	256	2	US-08-485-862B-51	Sequence 51, Appl
30	33	55.0	256	3	US-08-487-077A-51	Sequence 51, Appl
31	33	55.0	256	3	US-08-485-863A-51	Sequence 51, Appl
32	33	55.0	256	4	US-08-485-049D-51	Sequence 51, Appl
33	33	55.0	257	3	US-08-787-739-51	Sequence 87, Appl
34	33	55.0	377	3	US-08-787-739-87	Sequence 87, Appl
35	33	55.0	422	3	US-08-335-469-2	Sequence 2, Appl
36	33	55.0	449	2	US-08-839-008-9	Sequence 2, Appl
37	33	55.0	459	2	US-08-481-658B-2	Sequence 2, Appl
38	33	55.0	459	2	US-08-477-504A-2	Sequence 2, Appl
39	33	55.0	459	2	US-08-486-756A-2	Sequence 2, Appl
40	33	55.0	459	2	US-08-486-862B-2	Sequence 2, Appl
41	33	55.0	459	3	US-08-787-739-2	Sequence 2, Appl
42	33	55.0	459	3	US-08-487-077A-2	Sequence 2, Appl
43	33	55.0	459	3	US-08-485-863A-2	Sequence 2, Appl
44	33	55.0	459	4	US-08-485-049D-2	Sequence 2, Appl
45	33	55.0	1251	5	PCR-US95-02251-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-416-478A-6
Sequence 6, Application US/08416478A
Patent No. 5773578
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: New Proteins Produced By Human Lymphocytes, DNA Sequences Encoding These Proteins And Their Pharmaceutical And Biological Uses
TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,478A
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND=1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-478A-6

Query Match 65.0%; Score 39; DB 1; Length 380;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LOVLGQLPPT 12
| | | | |
Db 257 LTVLGLPPT 267

RESULT 2

US-08-474-988B-6
Sequence 6, Application US/08474988B
Patent No. 5874250

GENERAL INFORMATION:

APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,988B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/416,478
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644

FILING DATE: 08-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-474-988B-6

Query Match

Best Local Similarity 65.0%; Score 39; DB 2; Length 380;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LOVLGQLPPT 12
| | | | |
Db 257 LTVLGLPPT 267

RESULT 3

US-08-394-442B-6
Sequence 6, Application US/08394442B
Patent No. 5976877

GENERAL INFORMATION:

APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: New Proteins Produced By Human
LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND

TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,442B
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: HERCEND-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-394-442B-6

Query Match 65.0%; Score 39; DB 2; Length 380;
Best Local Similarity 72.7%; Pred. No. 13;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LOVLGQLPPT 12
| | | | |
Db 257 LTVLGLPPT 267

RESULT 4

US-08-416-478A-7
Sequence 7, Application US/08416478A
Patent No. 5773578

GENERAL INFORMATION:

APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: New Proteins Produced By Human
LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,478A
FILING DATE: 04-APR-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.

STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,478A
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 471 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /note="Hydrogen is present at the
OTHER INFORMATION: N-terminus"
US-08-416-478A-2

Query Match 65.0%; Score 39; DB 1; Length 471;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LQVIGLQLPPT 12
DB 230 LTVIGLEPPPT 240

RESULT 8
US-08-474-988B-2
Sequence 2, Application US/08474988B
Patent No. 5874250
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,988B
FILING DATE: 07-JUN-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/416,478
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 471 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /note="Hydrogen is present at the
OTHER INFORMATION: N-terminus"
US-08-474-988B-2

Query Match 65.0%; Score 39; DB 2; Length 471;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LQVIGLQLPPT 12
DB 230 LTVIGLEPPPT 240

RESULT 9
US-08-394-442B-2
Sequence 2, Application US/08394442B
Patent No. 5976877
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
APPLICANT: Triebel, Frederic
TITLE OF INVENTION: New Proteins Produced By Human
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,442B
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: HERCEND-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 471 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /note="Hydrogen is present at the
OTHER INFORMATION: N-terminus"
US-08-394-442B-2

Query Match 65.0%; Score 39; DB 2; Length 471;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 LOVLGLQLEPTP 12
| | | | | | | | | |
Db 230 LTVIGLEPPT 240

RESULT 10

US-08-737-271-1
Sequence 1, Application US/08737271
Patent No. 5955300
GENERAL INFORMATION:
APPLICANT: FAURE, Florence
APPLICANT: HERCEND, Thierry
APPLICANT: HUARD, Bertrand
APPLICANT: TRIEBEL, Frederic
TITLE OF INVENTION: SOLUBLE POLYPEPTIDE FRACTIONS OF THE
TITLE OF INVENTION: LAG-3 PROTEIN; PRODUCTION METHOD, THERAPEUTIC COMPOSITION,
NUMBER OF SEQUENCES: 8
TITLE OF INVENTION: ANTI-IDIOTYPE ANTIBODIES
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,271
FILING DATE: 24-DEC-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/05643
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR95/00593
FILING DATE: 05-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: BET 96/844
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-737-271-1

Query Match 65.0%; Score 39; DB 2; Length 476;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 LOVLGLQLEPTP 12
| | | | | | | | | |
Db 235 LTVIGLEPPT 245

RESULT 11

US-09-058-555-1
Sequence 1, Application US/09058555
Patent No. 6143273
GENERAL INFORMATION:
APPLICANT: FAURE, Florence
APPLICANT: HERCEND, Thierry
APPLICANT: HUARD, Bertrand
APPLICANT: TRIEBEL, Frederic
TITLE OF INVENTION: SOLUBLE POLYPEPTIDE FRACTIONS OF THE
TITLE OF INVENTION: LAG-3 PROTEIN; PRODUCTION METHOD, THERAPEUTIC COMPOSITION,
NUMBER OF SEQUENCES: 8
TITLE OF INVENTION: ANTI-IDIOTYPE ANTIBODIES
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/058,555
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/737,271
FILING DATE: 24-DEC-1996
APPLICATION NUMBER: FR 94/05643
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR95/00593
FILING DATE: 05-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: BET 96/844
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-058-555-1

Query Match 65.0%; Score 39; DB 4; Length 476;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 LOVLGLQLEPTP 12
| | | | | | | | | |
Db 235 LTVIGLEPPT 245

RESULT 12

US-08-416-478A-9

Sequence 9, Application US/08416478A
Patent No. 5773578

GENERAL INFORMATION:

APPLICANT: Hercend, Thierry
APPLICANT: Tiebel, Frederic
TITLE OF INVENTION: New Proteins Produced By Human
TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416/478A
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-416-478A-9
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-416-478A-9
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 65.0%; Score 39; DB 1; Length 498;
Best Local Similarity 72.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 257 LTVGLGLEPPT 267

US-08-474-988B-9
Sequence 9, Application US/08474988B
Patent No. 5874250

GENERAL INFORMATION:

APPLICANT: Hercend, Thierry
APPLICANT: Tiebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington

US-08-474-988B-9
Sequence 9, Application US/08474988B
Patent No. 5874250
APPLICANT: Hercend, Thierry
APPLICANT: Tiebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,988B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-474-988B-9
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-474-988B-9
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 65.0%; Score 39; DB 2; Length 498;
Best Local Similarity 72.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 257 LTVGLGLEPPT 267

US-08-394-442B-9
Sequence 9, Application US/08394442B
Patent No. 5976877

GENERAL INFORMATION:

APPLICANT: Hercend, Thierry
APPLICANT: Tiebel, Frederic
TITLE OF INVENTION: New Proteins Produced By Human
TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,442B
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

US-08-394-442B-9
Sequence 9, Application US/08394442B
Patent No. 5976877
APPLICANT: Hercend, Thierry
APPLICANT: Tiebel, Frederic
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

Tue Jun 5 07:09:35 2001

us-09-284-107-28.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:45 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-28

Perfect score: 60

Sequence: 1 ELQVIGLQPTP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	344	2 A41357	Fc gamma (IgG) rec
2	60	100.0	374	1 A39878	Fc gamma (IgG) rec
3	40	66.7	280	2 I55577	Fc gamma (IgG) rec
4	39	65.0	498	2 S11246	LAG-3 protein prec
5	38	63.3	174	2 B34421	male-enhanced anti
6	38	63.3	174	2 S39790	male-enhanced anti
7	38	63.3	185	2 A34421	male-enhanced anti
8	38	63.3	317	2 B82577	hypothetical prote
9	38	63.3	425	2 H64047	threonine synthase
10	37	61.7	134	2 T32913	hypothetical prote
11	37	61.7	226	2 E71478	probable phosphogl
12	37	61.7	253	2 C75611	transcription regu
13	37	61.7	424	2 T33839	hypothetical prote
14	37	61.7	3212	2 T24692	hypothetical prote
15	36	60.0	144	2 S54130	hypothetical 16.34
16	36	60.0	145	2 G70384	hypothetical prote
17	36	60.0	174	2 S59848	ribosomal protein
18	36	60.0	180	2 S56056	ribosomal protein
19	36	60.0	365	2 T36806	probable alcohol d
20	36	60.0	379	2 S42529	opaque-2-related p
21	36	60.0	408	2 S42493	opaque-2 protein
22	36	60.0	419	2 S56073	opaque-2 protein
23	36	60.0	437	2 A34800	opaque-2 protein
24	36	60.0	460	2 S06022	regulatory protein
25	36	60.0	572	2 T20764	hypothetical prote
26	36	60.0	596	2 A45195	adenylyl cyclase ty
27	36	60.0	609	2 T14759	adenylyl cyclase ty
28	36	60.0	1184	2 A42904	adenylyl cyclase ty
29	36	60.0	1223	2 S29717	adenylyl cyclase

30	36	60.0	1264	2 S41603	type V adenylyl cy
31	35	58.3	208	2 T13515	hypothetical prote
32	35	58.3	215	2 T47380	hypothetical prote
33	35	58.3	233	2 C72105	conserved hypotnet
34	35	58.3	273	2 C33282	DNA-binding protei
35	35	58.3	342	2 A83263	dihydroorotate deh
36	35	58.3	351	2 C64646	dihydroorotate deh
37	35	58.3	351	2 E71935	dihydroorotate deh
38	35	58.3	388	2 D75496	aspartate transami
39	35	58.3	397	2 T06183	serpin - barley
40	35	58.3	399	1 DXBHZ	protein 24 - Barle
41	35	58.3	410	2 T37078	probable Na+/H+ an
42	35	58.3	448	2 C56711	casein kinase I (E
43	35	58.3	459	2 F71257	hypothetical prote
44	35	58.3	550	1 S14048	RNA helicase dbp2
45	35	58.3	553	2 S06345	cell fusion glycop

ALIGNMENTS

RESULT 1
Fc gamma (IgG) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000
A:Accession: A41357; S03019
R:Allen, J.M.; Seed, B.
Science 243, 378-381, 1989
A:Title: Isolation and expression of functional high-affinity Fc receptor complementa
A:Reference number: A41357; MUID:89100284
A:Accession: A41357
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: GB:X14355; GB:M21090; NID:g31333; PIDN:CAA32536.1; PID:g31334
R:Allen, J.M.; Seed, B.
Nucleic Acids Res. 16, 11824, 1988
A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (A:Reference number: S03018; MUID:89098339
A:Accession: S03019
A:Molecule type: mRNA
A:Residues: 1-344 <ALL>
A:Cross-references: EMBL:X14355; NID:g31333; PIDN:CAA32536.1; PID:g31334
A>Note: the authors translated the codon ACT for residue 25 as Ala
C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
F:117-170/Domain: immunoglobulin homology <IMM>
Query Match 100.0%; Score 60; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ELQVIGLQPTP 12
Db 277 ELQVIGLQPTP 288
RESULT 2
Fc gamma (IgG) receptor I-A (high affinity) precursor - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
A:Accession: A39878; I70304; B41357; S03018; I57525
R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
J. Biol. Chem. 266, 13449-13455, 1991
A:Title: Gene organization of the human high affinity receptor for IgG, Fc gamma RI (CD
A:Reference number: A39878; MUID:91302383
A:Accession: A39878
A:Molecule type: DNA
A:Residues: 1-374 <VAN>

A:Cross-references: GB:M63830; GB:M63835; NID:g180277; PIDN:AAA35678.1; PID:g180279
 R:Porres, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: 155577; MUID:93055454
 A:Accession: 170304
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:L03418; NID:g184840; PIDN:AAA36049.1; PID:g292169
 R:Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: A41357; MUID:89100284
 A:Accession: B41357
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'T', '339-374' <AL2>
 R:Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR)
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'T', '339-374' <AL2>
 A:Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R:Perez, C.; Wietzerbin, J.; Benech, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A:Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int
 ism.
 A:Reference number: 157525; MUID:93204964
 A:Accession: 157525
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RE2>
 A:Cross-references: GB:S57204; NID:g298692; PIDN:AAD13887.1; PID:g4261587
 C:Genetics:
 A:Gene: GDB:FCGR1A; CD64
 A:Cross-references: GDB:135911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-292/Domain: extracellular #status predicted <EXT>
 F:117-170/Domain: immunoglobulin homology <IMM>
 F:293-313/Domain: transmembrane #status predicted <TM>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent) #status predicte

Query Match 100.0%; Score 60; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELQVIGLQLEPTP 12
 DB 277 ELQVIGLQLEPTP 288

RESULT 3
 155577
 Fc gamma (19g) receptor I-B splice form 1 precursor - human
 N:Alternate names: CD64
 C:Contains: Fc-gamma (19g) receptor I-B splice form 2
 C:Species: Homo sapiens (man)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
 C:Accession: 155577; 170303
 R:Porres, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: 155577; MUID:93055454
 A:Accession: 155577
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA

A:Residues: 1-280 <RES>
 A:Cross-references: GB:L03419; NID:g182460; PIDN:AAA35825.1; PID:g292023
 A:Note: splice form B1
 A:Accession: 170303
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-10, 103-153, 'A', '155-280' <RE2>
 A:Cross-references: GB:L03420; NID:g182461; PIDN:AAA35826.1; PID:g292024
 A:Experimental source: mononuclear cells
 A:Note: splice form B2
 C:Comment: This receptor does not bind monomeric IgG with high affinity.
 C:Genetics:
 A:Gene: GDB:FCGR1B; CD64
 A:Cross-references: GDB:135923; OMIM:601502
 A:Map position: 1p12-1p12
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembran
 F:117-170/Domain: immunoglobulin homology <IMM>

Query Match 66.7%; Score 40; DB 2; Length 280;
 Best Local Similarity 88.9%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 VLGIQLPTP 12
 DB 186 VKGIQLPTP 194

RESULT 4
 S11246
 LAG-3 protein precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 04-Mar-2000
 C:Accession: S11246
 R:Triebel, F.; Jitsukawa, S.; Balxeras, E.; Roman-Roman, S.; Genevee, C.; Viegas-Pegu
 J. Exp. Med. 171, 1393-1405, 1990
 A:Title: LAG-3, a novel lymphocyte activation gene closely related to CD4.
 A:Reference number: S11246; MUID:90237736
 A:Accession: S11246
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-498 <TR1>
 A:Cross-references: EMBL:X51985
 A:Note: the author translated the codon CCA for residue 388 as Arg
 C:Genetics:
 A:Gene: GDB:LAG3
 A:Cross-references: GDB:127449; OMIM:153337
 A:Map position: 12p13.3-12p13.3
 C:Superfamily: human LAG-3 protein
 C:Keywords: transmembrane protein

Query Match 65.0%; Score 39; DB 2; Length 498;
 Best Local Similarity 72.7%; Pred. No. 19;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LQVIGLQLEPTP 12
 DB 257 LTVIGLEPTP 267

RESULT 5
 B34421
 male-enhanced antigen - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 15-Jun-1990 #sequence_revision 09-Oct-1992 #text_change 05-Nov-1999
 C:Accession: B34421
 R:Lau, Y.F.C.; Chan, K.; Sparkes, R.
 Proc. Natl. Acad. Sci. U.S.A. 86, 8462-8466, 1989
 A:Title: Male-enhanced antigen gene is phylogenetically conserved and expressed at la
 A:Reference number: A34421; MUID:90046817
 A:Accession: B34421

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-174 <LAN>
A:Cross-references: GB:M27938; NID:g1919115; PIDN:AAA39519.1; PID:g1919116
A:Note: the authors translated the codon GCA for residue 31 as Ser, and ACA for residue

Query Match 63.3%; Score 38; DB 2; Length 174;
Best Local Similarity 63.6%; Pred. No. 9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LOVLGLOLPTP 12
Db 89 IQALGLHLPDP 99

RESULT 6
S39790
male-enhanced antigen - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-May-1994 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000
C:Accession: S39790
R:Kondo, M.; Sato, S.; Suto, S.
Biochim. Biophys. Acta 1216, 483-486, 1993
A:Title: Cloning and sequence analysis of cDNA encoding the bovine testis-derived male-
A:Reference number: S39790; MUID:94092744
A:Accession: S39790
A:Molecule type: mRNA
A:Residues: 1-174 <KON>
A:Cross-references: EMBL:D17340; NID:g560052; PIDN:BAA04158.1; PID:g560053
A:Note: It is uncertain whether Met-1 or Met-11 is the initiator

Query Match 63.3%; Score 38; DB 2; Length 174;
Best Local Similarity 63.6%; Pred. No. 9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LOVLGLOLPTP 12
Db 89 IQALGLHLPDP 99

RESULT 7
A34421
male-enhanced antigen - human
C:Species: Homo sapiens (man)
C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
C:Accession: A34421
R:Law, Y.F.C.; Chan, K.; Sparkes, R.
Proc. Natl. Acad. Sci. U.S.A. 86, 8462-8466, 1989
A:Title: Male-enhanced antigen gene is phylogenetically conserved and expressed at late
A:Reference number: A34421; MUID:90046817
A:Accession: A34421
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-185 <LAN>
A:Cross-references: GB:M27937; NID:g187507; PIDN:AAA36208.1; PID:g307182

Query Match 63.3%; Score 38; DB 2; Length 185;
Best Local Similarity 63.6%; Pred. No. 9.6;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LOVLGLOLPTP 12
Db 100 IQALGLHLPDP 110

RESULT 8
B82577
hypothetical protein XF2283 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82577
R:anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below

A:Accession: B82577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-317 <STM>

A:Cross-references: GB:AE004040; GB:AE003849; NID:g9107437; PIDN:AAF85082.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvares, R.
Briotes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carer
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2283

Query Match 63.3%; Score 38; DB 2; Length 317;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 ELQVGLQPLP 12
Db 182 ELRIGMDGPTP 193

RESULT 9
H64047
threonine synthase (EC 4.2.99.2) - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 16-Jul-1999
C:Accession: H64047
R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: H64047
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-425 <TIGR>
A:Cross-references: GB:U32694; GB:L42023; NID:g1573035; PIDN:AAC21765.1; PID:g1573038
C:Superfamily: threonine synthase
C:Keywords: carbon-oxygen lyase; phosphoprotein; pyridoxal phosphate; threonine biosy
F:105/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 63.3%; Score 38; DB 2; Length 425;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 QVGLQPLP 12
Db 386 RILGILPLP 395

RESULT 10
T32913

hypothetical protein C54G6.4 - *Caenorhabditis elegans* (fragment)

C:Species: *Caenorhabditis elegans*

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32913

R:Pin-Mollam, A.; Graves, T.; Ozersky, P.

A:Description: The sequence of *C. elegans* cosmid C54G6.

A:Reference number: 221245

A:Accession: T32913

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-134 <PIN>

A:Cross-references: EMBL:AF043698; PIDN:AA097562.1; GSPDB:GN00019; CESP:C54G6.4

A:Experimental source: strain Bristol N2; clone C54G6

C:Genetics:

A:Gene: CESP:C54G6.4

A:Map position: 1

A:introns: 36/3

Query Match 61.7% Score 37; DB 2; Length 134;

Best Local Similarity 70.0%; Pred. No. 10; Mismatches 2; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QVIGLQLPPT 12

DB 98 QQLGVQLPVP 107

RESULT 11

E71478

probable phosphoglycerate mutase - *Chlamydia trachomatis* (serotype D, strain UM3/Cx)

C:Species: *Chlamydia trachomatis*

C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 16-Jul-1999

C:Accession: E71478

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*

A:Reference number: A71570; MUID:99000809

A:Accession: E71478

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-226 <ARN>

A:Cross-references: GB:AE001343; GB:AE001273; NID:93329176; PIDN:AA068317.1; PID:9332917

A:Experimental source: serotype D, strain UM-3/Cx

C:Genetics:

A:Gene: pm

C:Superfamily: phosphoglycerate mutase; phosphoglycerate mutase homology

Query Match 61.7% Score 37; DB 2; Length 226;

Best Local Similarity 72.7%; Pred. No. 18; Mismatches 2; Indels 0; Gaps 0;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ELQVGLQLPT 11

DB 195 EEOVLSELPPT 205

RESULT 12

C75611

transcription regulator, *ICR* family - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: C75611

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896

A:Accession: C75611

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-253 <WH>

A:Cross-references: GB:AE001862; GB:AE001825; NID:96460468; PIDN:AA012212.1; PID:9646

A:Experimental source: strain R1

C:Genetics:

A:Gene: DRA0152

A:Map position: 2

C:Superfamily: acetate operon repressor

Query Match 61.7% Score 37; DB 2; Length 253;

Best Local Similarity 63.6%; Pred. No. 21; Mismatches 3; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LOVIGLQLPPT 12

DB 205 LAAIGVSLPPT 215

RESULT 13

T33839

hypothetical protein F54A5.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33839

R:Jones, K.; Graves, T.; Ozersky, P.; Wilson, R.

A:Description: The sequence of *C. elegans* cosmid F54A5.

A:Reference number: 221420

A:Accession: T33839

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-424 <CON>

A:Cross-references: EMBL:AF106584; PIDN:AA078220.1; GSPDB:GN00019; CESP:F54A5.1

A:Experimental source: strain Bristol N2; clone F54A5

C:Genetics:

A:Gene: CESP:F54A5.1

A:Map position: 1

A:introns: 154/1; 326/3

Query Match 61.7% Score 37; DB 2; Length 424;

Best Local Similarity 70.0%; Pred. No. 37; Mismatches 2; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QVIGLQLPPT 12

DB 388 QQLGVQLPVP 397

RESULT 14

T24692

hypothetical protein T08G11.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T24692

R:Dodson, R.

A:Title: The sequence of *C. elegans* cosmid T08G11.1

A:Reference number: 219925

A:Accession: T24692

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3212 <MIL>

A:Cross-references: EMBL:Z80220; PIDN:CA02304.1; GSPDB:GN00019; CESP:T08G11.1

A:Experimental source: clone T08G11

C:Genetics:

A:Gene: CESP:T08G11.1

A:Map position: 1

A:introns: 34/1; 149/3; 182/3; 527/2; 766/1; 1070/3; 1769/3; 1834/3; 2070/2; 2210/3;

Query Match 61.7% Score 37; DB 2; Length 3212;

Best Local Similarity 63.6%; Pred. No. 3; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LOVLGLQLPTP 12

DB 772 LMKLGISLPTP 782

RESULT 15

SS4130

hypothetical 16.34k protein - fowl adenovirus 1

C:Species: Avidenovirus gall (fowl adenovirus 1, CEL0)

C:Date: 08-Jul-1995 #sequence_revision 14-Feb-1997 #text_change 22-Oct-1999

C:Accession: SS4130

R:Hess, M.; Cuzange, A.; Chroboczek, J.; Ruigrok, R.; Jacrot, B.

Submitted to the EMBL Data Library, February 1995

A:Description: The sequence of the two fibers of an avian adenovirus (CEL0) and organise

A:Reference number: SS4125

A:Accession: SS4130

A:Molecule type: DNA

A:Residues: 1-144 <HES>

A:Cross-references: EMBL:X84724; NID:q780165; PIDN:CA59211.1; PID:q780172

R:Hess, M.; Cuzange, A.; Ruigrok, R.W.H.; Chroboczek, J.; Jacrot, B.

J. Mol. Biol. 252, 379-385, 1995

A:Title: The avian adenovirus penton: two fibres and one base.

A:Reference number: S59067; MUID:96025073

A:Contents: annotation

C:Superfamily: fowl adenovirus 1 hypothetical 16.34k protein

Query Match 60.0%; Score 36; DB 2; Length 144;

Best Local Similarity 85.7%; Pred.No. 17;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 GLQLPTP 12

DB 39 GLQMPPTP 45

Search completed: June 4, 2001, 12:15:46
Job time: 279 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:08 ; Search time 37.56 Seconds
(Without alignments)
10.944 Million cell updates/sec

Title: US-09-284-107-28

Sequence: 1 ELQVGLQLPTP 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	374	1	FCGL_HUMAN
2	39	65.0	525	1	P18314 homo sapien
3	38	63.3	113	1	LAG3_HUMAN
4	38	63.3	172	1	MEAL_PIG
5	38	63.3	174	1	MEAL_HUMAN
6	38	63.3	174	1	MEAL_BOVIN
7	38	63.3	174	1	MEAL_MOUSE
8	38	63.3	474	1	THRC_HABIN
9	38	60.0	174	1	P2X2_CAVPO
10	36	60.0	195	1	R120_YEAST
11	36	60.0	292	1	BID_MOUSE
12	36	60.0	453	1	Y161_HUMAN
13	36	60.0	572	1	OP2_MAIZE
14	36	60.0	572	1	NH25_CAEEL
15	36	60.0	1264	1	CVAS_RAT
16	36	60.0	1264	1	CVAS_RABIT
17	35	58.3	351	1	CVAS_CANFA
18	35	58.3	351	1	PYRD_HELPJ
19	35	58.3	399	1	PRYZ_HELPJ
20	35	58.3	447	1	KC13_HUMAN
21	35	58.3	448	1	KC13_RAT
22	35	58.3	550	1	DBP2_SCHPO
23	35	58.3	553	1	DBP2_SCHPO
24	34	56.7	176	1	VLGF_NDVI
25	34	56.7	176	1	VLGF_NDVI
26	34	56.7	328	1	RLIX_RAT
27	34	56.7	404	1	FCGL_MOUSE
28	34	56.7	463	1	ROK_HUMAN
29	34	56.7	463	1	ROK_RABIT
30	34	56.7	464	1	ROK_MOUSE
31	34	56.7	464	1	ROK_MOUSE
32	34	56.7	572	1	PTL_STAU
33	34	56.7	574	1	PTL_STACA

34	34	56.7	585	1	YAOH_SCHRO	010096 schizosacch
35	34	56.7	681	1	BRC2_HAIRO	001409 halocynthia
36	34	56.7	2787	1	TELL_YEAST	P28110 saccharomyc
37	33	55.0	109	1	YC44_METAP	050249 methanococc
38	33	55.0	203	1	VATD_CHLMU	09pk87 chlamydia m
39	33	55.0	203	1	VATD_CHLMU	084308 chlamydia t
40	33	55.0	213	1	YOUN_CAEEL	P34658 caenorhabdi
41	33	55.0	231	1	BID2_ECOLI	P77201 escherichia
42	33	55.0	249	1	Y768_METJA	058178 methanococc
43	33	55.0	296	1	SAPC_ECOLI	047624 escherichia
44	33	55.0	296	1	SAPC_ECOLI	P36669 salmonella
45	33	55.0	322	1	PMUC_SALTY	P24520 salmonella

ALIGNMENTS

RESULT 1
FCGL_HUMAN STANDARD: PRT: 374 AA.
AC P12314; P12315;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (TGG FC RECEPTOR I) (CD64).
GN FCGR1A OR FCGR1 OR FCGL OR IGFRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89098339; PubMed=2974947;
RA Allen J.M., Seed B.;
RT "Nucleotide sequence of three cDNAs for the human high affinity FC receptor (FcRI).";
RL Nucleic Acids Res. 16:11824-11824(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89100284; PubMed=2911749;
RA Allen J.M., Seed B.;
RT "Isolation and expression of functional high-affinity FC receptor complementary DNAs.";
RL Science 243:378-381(1989).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH AFFINITY RECEPTOR.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD64 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd64.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC EMBL: X14356; CA32537.1;
DR EMBL: X14355; CA32536.1;
DR PIR: S03018; S03018.
DR PIR: S03019; S03019.
DR PIR: A41357; A41357.
DR PIR: B41357; B41357.
DR HSSP: P12319; IALT.
DR MIM: 146760;
DR InterPro: IPR003006;
DR Pfam: PF00047; Ig; 3.

KM IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain; Alternative splicing; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 374
 FT DOMAIN 22 292
 FT TRANSMEM 293 313
 FT DOMAIN 314 374
 FT DOMAIN 36 92
 FT DOMAIN 117 175
 FT DOMAIN 205 267
 FT DISULFID 43 85
 FT DISULFID 124 168
 FT DISULFID 212 260
 FT CARBOHYD 59 59
 FT CARBOHYD 78 78
 FT CARBOHYD 152 152
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT VARSPLIC 333 374
 FT VARIANT 26 26
 FT VARIANT 59 59
 FT CONFLICT 25 25
 FT SEQUENCE 374 AA; 42605 MW; 2C2AAB103BC16E6 CMC64;
 Query Match
 Best Local Similarity 100.0%; Score 60; DB 1; Length 374;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ELOVGLQLEPT 12
 DB 277 ELOVGLQLEPT 288
 RESULT 2
 LAG3_HUMAN STANDARD; PRT; 525 AA.
 AC P18627;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LYMPHOCYTE ACTIVATION GENE-3 PROTEIN PRECURSOR (LAG-3) (FDC PROTEIN).
 GN LAG3 OR FDC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90237736; PubMed-1692078;
 RA Triebel F., Jitsukawa S., Balxeras E., Roman-Roman S., Genevee C.,
 RT Viegas-Pequignot E., Hercend T.;
 RL J. Exp. Med. 171:1393-1405(1990).
 RN (2).
 RP REVISIONS TO C-TERMINUS.
 RA Triebel F.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RP CHARACTERIZATION.
 RX MEDLINE-92364535; PubMed-1380059;
 RA Balxeras E., Huard B., Miossec C., Jitsukawa S., Marlin M.,
 RT Hercend T., Auffray C., Miossec C., Triebel F., Platzer-Touneau D.;
 RL "Characterization of the lymphocyte activation gene 3-encoded
 antigens".
 RT J. Exp. Med. 176:327-337(1992).

CC -1- FUNCTION: INVOLVED IN LYMPHOCYTE ACTIVATION. BINDS TO HLA CLASS-II
 CC ANTIGENS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: ON CELL SURFACE OF ACTIVATED NK AND
 CC T-LYMPHOCYTES.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
 CC ONE V-LIKE AND 3 C2-LIKE DOMAINS. SIMILAR TO CD4.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X51985; CA36243.1; -
 CC EMBL: A21353; CA01547.1; ALT-SEQ.
 CC PIR: S11246; S11246.
 CC MIM: 153337; -
 CC InterPro: IPR003006;
 CC Pfam: PF00047; 1g; 3.
 KW Signal; Transmembrane; Glycoprotein; Immunoglobulin domain; Repeat.
 FT SIGNAL 1 28
 FT CHAIN 29 525
 FT DOMAIN 29 450
 FT TRANSMEM 451 471
 FT DOMAIN 472 525
 FT DOMAIN 37 167
 FT DOMAIN 182 248
 FT DOMAIN 275 340
 FT DOMAIN 362 419
 FT DOMAIN 505 523
 FT DISULFID 44 160
 FT DISULFID 189 241
 FT DISULFID 282 333
 FT DISULFID 369 412
 FT CARBOHYD 188 188
 FT CARBOHYD 250 250
 FT CARBOHYD 256 256
 FT CARBOHYD 343 343
 FT SEQUENCE 525 AA; 57495 MW; C447DB8B0E9E2733 CMC64;
 Query Match
 Best Local Similarity 65.0%; Score 39; DB 1; Length 525;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 2 LQVGLGLEPT 12
 DB 257 LQVGLGLEPT 267
 RESULT 3
 MEAL_PIG STANDARD; PRT; 113 AA.
 AC Q95313;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MALE-ENHANCED ANTIGEN-1 (MEA-1) (FRAGMENT).
 GN MEAL OR MEA.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OX NCBI_TaxID=9623;
 RP SEQUENCE FROM N.A.
 RX Tissue-Small Intestine;
 RA Winteroe A.K., Fredholm M., Davies W.;
 RT Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RL -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
 CC TESTIS DEVELOPMENT.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 281165; CAB03550.1; -
 CC Spermatoogenesis; Developmental protein.
 CC NON_TER 113
 CC SEQUENCE 113 AA; 12083 MW; 1F52DC82FDA95B6 CRC64;

Query Match 63.3%; Score 38; DB 1; Length 113;
 Best Local Similarity 63.6%; Pred. No. 2.4;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LOVLGLOLPTP 12
 Db 89 IOALGLHLDP 99

RESULT 4
 MEAL_HUMAN
 ID MEAL_HUMAN STANDARD; PRT; 172 AA.
 AC Q16626;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MALE-ENHANCED ANTIGEN-1 (MEA-1).
 GN MEAL OR MEA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=90046817; Pubmed-2813404;
 RX Lau Y.-F.C., Chan K., Sparkes R.S.;
 RT "Male-enhanced antigen gene is phylogenetically conserved and
 RT expressed at late stages of spermatogenesis".
 RL Proc. Natl. Acad. Sci. U.S.A. 86:8462-8466(1989).
 CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
 CC TESTIS DEVELOPMENT.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS.
 CC -1- CAUTION: WAS ORIGINALLY THOUGHT TO BE THE H-Y ANTIGEN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-11 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M27937; AAA36208.1; ALT_INIT.
 CC EMBL; L10400; AAA36209.1; ALT_INIT.
 CC MIM; 143170; -
 CC Spermatoogenesis; Developmental protein.
 CC SEQUENCE 172 AA; 18544 MW; 9342AFAD2C96FF25 CRC64;

Query Match 63.3%; Score 38; DB 1; Length 172;
 Best Local Similarity 63.6%; Pred. No. 3.7;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LOVLGLOLPTP 12
 Db 87 IOALGLHLDP 97

RESULT 5
 MEAL_BOVIN
 ID MEAL_BOVIN STANDARD; PRT; 174 AA.
 AC Q29407;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MALE-ENHANCED ANTIGEN-1 (MEA-1).
 GN MEAL OR MEA.
 OS Bos taurus (Bovinae).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=94092744; Pubmed-8268232;
 RA Kondo M., Sato S., Sutoh S.;
 RT "Cloning and sequence analysis of cDNA encoding the bovine testis-
 RT derived male-enhanced antigen (Mea)."
 RL Biochim. Biophys. Acta 1216:483-486(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97063437; Pubmed-8907304;
 RA Kondo M., Terouchi S., Tsukasa N., Sato S., Ishida N., Sutoh S.;
 RT "Genomic sequence analysis of the bovine male-enhanced antigen-1
 RT (Mea-1) and differential localization of its transcripts and products
 RT during spermatogenesis".
 RL DNA Seq. 6:75-85(1996).
 CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
 CC TESTIS DEVELOPMENT.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
 CC FOUND IN PRIMARY AND SECONDARY SPERMATOCYTES, AND SPERMATIDS, BUT
 CC THE PROTEIN ITSELF IS ONLY DETECTED IN SPERMATIDS. NO EXPRESSION
 CC IN LEYDIG CELLS, SPERMATOGONIA, OR SPERM. VERY WEAK EXPRESSION IN
 CC THE HEART, KIDNEY, SPLEEN, THYMUS, AND OVARY.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE STAGES FROM 8-CELL EMBRYOS
 CC TO HATCHED BLASTOCYSTS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-11 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D17340; BAA04158.1; -
 CC EMBL; D30811; BAA06488.1; -
 CC Spermatoogenesis; Developmental protein.
 CC SEQUENCE 174 AA; 18751 MW; E0E7CE990CB3530 CRC64;

Query Match 63.3%; Score 38; DB 1; Length 174;
 Best Local Similarity 63.6%; Pred. No. 3.8;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LOVLGLOLPTP 12
 Db 89 IOALGLHLDP 99

RESULT 6
 MEAL_MOUSE
 ID MEAL_MOUSE STANDARD; PRT; 174 AA.
 AC Q64327;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MALE-ENHANCED ANTIGEN-1 (MEA-1).
 GN MEAL OR MEA.
 OS Mus musculus (Muridae).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Muridae; Muridae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=94092744; Pubmed-8268232;
 RA Kondo M., Sato S., Sutoh S.;
 RT "Cloning and sequence analysis of cDNA encoding the bovine testis-
 RT derived male-enhanced antigen (Mea)."
 RL Biochim. Biophys. Acta 1216:483-486(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97063437; Pubmed-8907304;
 RA Kondo M., Terouchi S., Tsukasa N., Sato S., Ishida N., Sutoh S.;
 RT "Genomic sequence analysis of the bovine male-enhanced antigen-1
 RT (Mea-1) and differential localization of its transcripts and products
 RT during spermatogenesis".
 RL DNA Seq. 6:75-85(1996).
 CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
 CC TESTIS DEVELOPMENT.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
 CC FOUND IN PRIMARY AND SECONDARY SPERMATOCYTES, AND SPERMATIDS, BUT
 CC THE PROTEIN ITSELF IS ONLY DETECTED IN SPERMATIDS. NO EXPRESSION
 CC IN LEYDIG CELLS, SPERMATOGONIA, OR SPERM. VERY WEAK EXPRESSION IN
 CC THE HEART, KIDNEY, SPLEEN, THYMUS, AND OVARY.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE STAGES FROM 8-CELL EMBRYOS
 CC TO HATCHED BLASTOCYSTS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-11 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D17340; BAA04158.1; -
 CC EMBL; D30811; BAA06488.1; -
 CC Spermatoogenesis; Developmental protein.
 CC SEQUENCE 174 AA; 18751 MW; E0E7CE990CB3530 CRC64;

DE MALE-ENHANCED ANTIGEN-1 (ME-1).
 GN MEAL OR MEA.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBL_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90046817; PubMed=2813404;
 RX Lau Y.-F.C., Chan K., Sparkes R.S.;
 RT "Male-enhanced antigen gene is phylogenetically conserved and
 RL expressed at late stages of spermatogenesis.";
 RT Proc. Natl. Acad. Sci. U.S.A. 86:8462-8466(1989).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=94092744; PubMed=8268232;
 RA Kondo M., Sato S., Sutou S.;
 RT "Cloning and sequence analysis of cDNA encoding the bovine testis-
 RL derived male-enhanced antigen (Mea)."
 CC Biochim. Biophys. Acta 1216:483-486(1993).
 CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
 CC TESTIS DEVELOPMENT.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
 CC FOUND IN PRIMARY AND SECONDARY SPERMATOCYTES, AND SPERMATIDS, BUT
 CC IN LEYDIG CELLS, SPERMATOGENIA, OR SPERM. VERY WEAK EXPRESSION IN
 CC THE HEART, KIDNEY, SPLEEN, THYMUS, AND OVARY.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AS EARLY AS DAY 6 POST-PARTUM, WITH
 CC HIGHER EXPRESSION IN THE ADULT TESTIS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-11 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M27938; AAA39519.1;
 DR EMBL: L10401; AAA39520.1;
 DR EMBL: D17341; BAA04159.1;
 DR MGD: MGI:96957; Mea1
 KW Spermatogenesis; Developmental protein.
 SQ SEQUENCE 174 AA; 18584 MW; C3B16361635B176F CRC64;

Query Match 63.3%; Score 38; DB 1; Length 174;
 Best Local Similarity 63.6%; Pred. No. 3.8;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LOVIGLOLPP 12
 DB 89 IQAGLHLPDP 99

RESULT 7
 THRC HAEIN STANDARD; PRT; 425 AA.
 AC P44503;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE THREONINE SYNTHASE (EC 4.2.99.2).
 GN THRC OR HI0087.
 OS Haemophilus influenzae.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 CC NCBL_TaxID=727;
 CC [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleisemann A.R., Adams M.D., White O., Clayton R.A., Kirness E.F.,
 RA Kerlayage R.D., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae RD.";
 RL Science 269:496-512(1995).
 CC -1- CATALYTIC ACTIVITY: O-PHOSPHO-L-HOMOSERINE + H(2)O -> L-THREONINE +
 CC ORTHOPHOSPHATE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: THREONINE BIOSYNTHESIS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32694; AAC21765.1;
 DR TIGR: HI0087;
 DR InterPro: IPR000634;
 DR InterPro: IPR001926;
 DR Pfam: PF00291; S-T dehydratase; 1
 DR PROSITE: PS00165; DEHYDRATASE, SER, THR, 1.
 KW Threonine biosynthesis; Lyase; Pyridoxal phosphate.
 FT BINDING 105 105 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 425 AA; 46682 MW; 2CC222DD201F09ED CRC64;

Query Match 63.3%; Score 38; DB 1; Length 425;
 Best Local Similarity 60.0%; Pred. No. 9.8;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 OVIGLOLPP 12
 DB 386 RILGLOLPP 395

RESULT 8
 P2X2_CAVPO STANDARD; PRT; 474 AA.
 ID P2X2_CAVPO
 AC 070397; Q70398;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE P2X PURINORECEPTOR 2 (ATP RECEPTOR) (P2X2) (PURINERGIC RECEPTOR).
 GN P2X2.
 OS Cavia porcellus (Guinea pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
 CC NCBL_TaxID=10141;
 CC [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-NIH 2; TISSUE=Organ of corti;
 RX MEDLINE=98346338; PubMed=9682808;
 RA Parker M.S., Laroque M.L., Campbell J.M., Bobbin R.P.,
 RA Deinelinger P.L.;
 RT "Novel variant of the P2X2 ATP receptor from the guinea pig organ of
 RT Corti.";
 RL Hear. Res. 121:62-70(1998).
 CC -1- FUNCTION: BINDING OF THIS LIGAND GATED ION CHANNEL TO ATP
 CC MEDIATES SYNAPTIC TRANSMISSION BETWEEN NEURONS AND FROM NEURONS TO
 CC SMOOTH MUSCLE.
 CC -1- SUBUNIT: HOMO- OR HETEROPOLYMERS.

CONFLICT CONFLICT SEQUENCE	144 231 453 AA;	144 231 49356 MM;	D -> A (IN REF. 2). K -> KR (IN REF. 2). 513A8AB8D5ABD999 CRC64

OY 2 LOVGLQJPTP 12
 DB 313 VQVVGLLPQP 323

RESULT 15

CVAS_RABIT STANDARD; PRT; 1264 AA.
 AC P40144;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ADENYLATE CYCLASE, TYPE V (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE)
 DE (ADENYLATE CYCLASE).
 GN ADCCY.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID:9986;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Myocardium;
 RA MEDLINE=94139935; PubMed=8307190;
 RA Wallach J., Droste M., Kluxen F.W., Pfeuffer T., Frank R.;
 RT "Molecular cloning and expression of a novel type V adenylyl cyclase
 from rabbit myocardium."
 RL FEBS Lett. 338:257-263(1994).
 CC -1- FUNCTION: THIS A MEMBRANE-BOUND, CA(2+)-INHIBITABLE ADENYLATE
 CYCLASE.
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.
 CC -1- ENZYME REGULATION: INHIBITION BY CA(2+) IN THE SUBMICROMOLAR
 CONCENTRATION RANGE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MYOCARDIAL TISSUE.
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: BELONGS TO ADENYLATE CYCLASE CLASS-4/GUANYLYL CYCLASE
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See: <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; Z29371; CA82562.1;
 CC DR HSSP; P19754; IAW.
 CC DR InterPro; IPR001054;
 CC DR Pfam; PF00211; guanylate_cyc; 2.
 CC DR PROSITE; PS00453; GUANYLATE_CYCLASES_1; 2.
 CC DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 2.
 CC KW Lyase; CAMP synthesis; Transmembrane; Glycoprotein; Duplication.
 CC FT DOMAIN 1 244 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 245 265 POTENTIAL.
 CC FT TRANSMEM 271 290 POTENTIAL.
 CC FT TRANSMEM 301 322 POTENTIAL.
 CC FT TRANSMEM 331 348 POTENTIAL.
 CC FT TRANSMEM 351 369 POTENTIAL.
 CC FT TRANSMEM 377 398 POTENTIAL.
 CC FT DOMAIN 399 765 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 766 786 POTENTIAL.
 CC FT TRANSMEM 797 816 POTENTIAL.
 CC FT TRANSMEM 839 859 POTENTIAL.
 CC FT DOMAIN 860 912 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 913 933 POTENTIAL.
 CC FT TRANSMEM 938 958 POTENTIAL.
 CC FT TRANSMEM 987 1006 POTENTIAL.
 CC FT DOMAIN 1007 1264 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 64 70 POLY-GIN.
 CC FT DOMAIN 79 82 POLY-ASP.
 CC FT DOMAIN 145 151 POLY-ALA.

FT CARBOHYD 873 873 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 975 975 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1264 AA; 139623 MW; 1787EB42A0C2FDF6 CRC64;

Query Match 60.0%; Score 36; DB 1; Length 1264;
 Best Local Similarity 63.6%; Pred. No. 75;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 2 LOVGLQJPTP 12
 DB 313 VQVVGLLPQP 325

Search completed: June 4, 2001, 12:24:09
 Job time: 571 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:23:30 ; Search time 107.68 Seconds
(without alignments)
13.062 Million cell updates/sec

Title: US-09-284-107-28
Perfect score: 60
Sequence: 1 ELQVIGLOLPTP 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.rodent:*
14: sp.vertebrate:*
15: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	374	4	092663
2	56	93.3	375	4	092495
3	40	66.7	188	4	092638
4	40	66.7	280	4	092637
5	40	66.7	349	6	09M270
6	40	66.7	502	2	09RGV6
7	39	65.0	1322	5	09N155
8	39	65.0	1322	5	09NAT0
9	38	63.3	247	5	09NAL2
10	38	63.3	317	2	09PB61
11	38	63.3	501	11	070399
12	38	63.3	649	4	09N8B8
13	37	61.7	134	5	044982
14	37	61.7	226	2	084727
15	37	61.7	253	2	09R201
16	37	61.7	424	5	09TYR0
17	37	61.7	3212	5	094010
18	36	60.0	144	14	064791
19	36	60.0	145	2	067106

20	36	60.0	163	5	090W7	09u0w7 leishmania
21	36	60.0	196	11	09JL76	09j1k6 rattus norv
22	36	60.0	196	11	09JL76	09j1k6 rattus norv
23	36	60.0	205	14	064763	064763 avian adeno
24	36	60.0	211	10	09SFX8	09s1x8 arabidopsis
25	36	60.0	265	10	09S966	09s966 zea mays (m
26	36	60.0	272	4	09S004	09s004 homo sapien
27	36	60.0	280	10	09XET7	09xet7 arena fatua
28	36	60.0	329	4	09S144	09s144 homo sapien
29	36	60.0	350	2	P82604	P82604 bacillus sp
30	36	60.0	363	3	09Y7D0	09y7d0 aspergillus
31	36	60.0	365	2	09S247	09s247 streptomyce
32	36	60.0	395	4	09Y308	09y398 homo sapien
33	36	60.0	408	10	039532	039532 coix laachry
34	36	60.0	419	10	007795	007795 sorghum bic
35	36	60.0	437	10	041835	041835 zea mays (m
36	36	60.0	492	5	09NJ96	09nj96 caenorhabdi
37	36	60.0	568	5	09NJ97	09nj97 caenorhabdi
38	36	60.0	609	4	09UG00	09ug00 homo sapien
39	36	60.0	639	4	09UJY5	09ujy5 homo sapien
40	36	60.0	1048	5	03VW5	03vw5 drosophila
41	36	60.0	1223	11	09QW3	09qw3 rattus sp.
42	36	60.0	1251	5	09NG79	09ng79 trichomonas
43	35	58.3	64	5	09VVG2	09vvg2 drosophila
44	35	58.3	115	2	09S1D8	09s1d8 yersinia pe
45	35	58.3	145	14	09WGF7	09wgf7 human papil

ALIGNMENTS

RESULT 1
092663 ID 092663 PRELIMINARY: PRT: 374 AA.
AC 092663;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE FC GAMMA RECEPTOR I.
GN A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=93055454; PubMed=1430234;
RA Porges A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RT Kimberly R.P.;
RT "Novel Fc gamma receptor I family gene products in human mononuclear
cells";
RL J. Clin. Invest. 90:2102-2109(1992).
DR EMBL: L03418; AAB36049.1;
DR HSSP: P12319; IALT.
DR INTERPRO: IPR003006;
DR PIRAM: PF00047; Ig; 3.
DR PRODOM: PD002534; -; 1.
SQ SEQUENCE 374 AA: 42632 MW: D33D59398CEEA699 CRC64;

Query Match 100.0%; Score 60; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELQVIGLOLPTP 12
DB 277 ELQVIGLOLPTP 288

RESULT 2
ID 092495 PRELIMINARY: PRT: 375 AA.

AC 092495;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last; sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last; annotation update)
DE FC GAMMA RECEPTOR TYPE I (FC GAMMA RIB=FC GAMMA RECEPTOR)
GN CD64 OR FC-GAMMA-RIB
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Benesh P.D., Sastry K.N., Iyer R.R., Eichbaum Q.G., Rayeh D.,
RA Ezekowitz A.B.,
RL Submitted (DEC-1992) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-224 FROM N.A.
RX MEDLINE=93018827; PubMed=1402657;
RA Benesh P.D., Sastry K., Iyer R.R., Eichbaum Q.G., Rayeh D.P.,
RA Ezekowitz R.A.,
RT "Definition of interferon gamma-response elements in a novel human Fc
gamma receptor gene (Fc gamma RIB) and characterization of the gene
structure."
RL J. Exp. Med. 176:1115-1123(1992).
DR EMBL; M91555; AAA58414.1; JOINED
DR EMBL; M91550; AAA58414.1; JOINED
DR EMBL; M91551; AAA58414.1; JOINED
DR EMBL; M91552; AAA58414.1; JOINED
DR EMBL; M91553; AAA58414.1; JOINED
DR EMBL; M91554; AAA58414.1; JOINED
DR EMBL; S45709; AAD13842.1; JOINED
DR EMBL; S45707; AAD13842.1; JOINED
DR EMBL; S45708; AAD13842.1; JOINED
DR EMBL; S45704; AAD13842.1; JOINED
DR HSSP; P12319; 1ALT
DR INTERPRO: IPR003006;
DR PFAM; PF00047; 19; 3;
DR PRODOM; PD002534; -1;
SO SEQUENCE 375 AA; 42881 MW; A84D464C70DD0F91 CRC64;

Query Match
Best Local Similarity 93.3%; Score 56; DB 4; Length 375;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ELQVIGLQLP 12
DB 278 ELQVIGLRLPT 289

RESULT 3
ID 092638 PRELIMINARY; PRT; 188 AA.
AC 092638;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last; sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last; annotation update)
DE FC GAMMA RECEPTOR I.
GN B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA MEDLINE=9305454; PubMed=1430234;
RA Forges A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RA Kimberly R.P.,
RT "Novel Fc gamma receptor I family gene products in human mononuclear
cells".

RL J. Clin. Invest. 90:2102-2109(1992).
DR EMBL; L03420; AAA35826.1; -
DR HSSP; P12319; 1ALT
DR INTERPRO: IPR003006;
DR PFAM; PF00047; 19; 1;
DR PRODOM; PD002534; -1;
SO SEQUENCE 188 AA; 22106 MW; E81517B33BEAB789 CRC64;

Query Match
Best Local Similarity 88.9%; Score 40; DB 4; Length 188;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 VLGLQLPT 12
DB .94 VKGLQLPT 102

RESULT 4
ID 092637 PRELIMINARY; PRT; 280 AA.
AC 092637;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last; sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last; annotation update)
DE FC GAMMA RECEPTOR I.
GN B1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=9305454; PubMed=1430234;
RA Forges A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
RA Kimberly R.P.,
RT "Novel Fc gamma receptor I family gene products in human mononuclear
cells".
RL J. Clin. Invest. 90:2102-2109(1992).
DR EMBL; L03419; AAA35825.1; -
DR HSSP; P12319; 1ALT
DR INTERPRO: IPR003006;
DR PFAM; PF00047; 19; 2;
DR PRODOM; PD002534; -1;
SO SEQUENCE 280 AA; 32232 MW; C6C6C45AB3D345C6 CRC64;

Query Match
Best Local Similarity 66.7%; Score 40; DB 4; Length 280;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 VLGLQLPT 12
DB 186 VKGLQLPT 194

RESULT 5
ID 09M270 PRELIMINARY; PRT; 349 AA.
AC 09M270;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last; sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last; annotation update)
DE FC GAMMA RECEPTOR I.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;
OX Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Li X., Wang A., Zhang G.;

RT "Molecular cloning and identification of full-length cDNA encoding
high affinity Fc receptor for bovine IgG (Fc gamma RI).";
RL Vet. Immunol Immunopathol. 75:151-159(2000).
DR EMBL; AF162866; AAF80477.1; -
KW Receptor.
SQ SEQUENCE 349 AA; 39608 MW; DOB77B2EF9408C02 CRC64;

Query Match 66.7%; Score 40; DB 6; Length 349;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ELOVIGLOLPTP 12
DB 277 ELPVIGLOLSTTP 288

RESULT 6
ID O9RGV6 PRELIMINARY; PRT; 502 AA.
AC O9RGV6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE TRHW ('PUTATIVE TRANSFER PROTEIN, PILUS FORMATION').
GN TRHW.
OS Salmomella typhi.
OC Plasmid R27.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Salmomella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99296679; PubMed-10366528;
RA Rooker M.M., Sherburne C., Lawley T.D., Taylor D.E.;
RT "Characterization of the Tra2 region of the IncH11 plasmid R27.";
RL Plasmid 41:226-239(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grobeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncH1 plasmid
from Salmomella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
DR EMBL; AF105019; AAD54026.1; -
DR EMBL; AF250878; AAF69846.1; -
KW Plasmid.
SQ SEQUENCE 502 AA; 56512 MW; F2FEF73D12B98FB6 CRC64;

Query Match 66.7%; Score 40; DB 2; Length 502;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ELOVIGLOLPTP 12
DB 42 QNOVLGAOLPVP 53

RESULT 7
ID O9NJS5 PRELIMINARY; PRT; 1322 AA.
AC O9NJS5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE SERINE PROTEASE 22D.
GN SP22D.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Culicidae; Anopheles.
OX NCBI_TaxID=7165;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-G3; TISSUE-HEMOLYMPH;
RX MEDLINE-20110889; PubMed-10646969;
RA Gorman M.J., Andreeva O.V., Paskewitz S.M.;
RT "Molecular characterization of five serine protease genes cloned from
Anopheles gambiae hemolymph.";
RL Insect Biochem. Mol. Biol. 30:35-46(2000).
DR EMBL; AF117751; AAD38337.3; -
KW Protease.
SQ SEQUENCE 1322 AA; 146811 MW; 2707110783A6B43 CRC64;

Query Match 65.0%; Score 39; DB 5; Length 1322;
Best Local Similarity 63.6%; Pred. No. 1e+02; 2; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LQVLGLOLPTP 12
DB 591 VOYIGLPAPTP 601

RESULT 8
ID O9NAT0 PRELIMINARY; PRT; 1322 AA.
AC O9NAT0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ADHESIVE SERINE PROTEASE.
GN SP22D.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Culicidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RA Danieli A., Loukeris T., Lagueux M., Mueller H.M., Richman A.,
RA Kafatos F.C.;
RT "A modular chitin-binding protease associated with hemocytes and
hemolymph in the mosquito Anopheles gambiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7136-7141(2000).
DR EMBL; AJ276428; CAB81934.1; -
KW Protease.
SQ SEQUENCE 1322 AA; 146794 MW; FB973C21CC5475B CRC64;

Query Match 65.0%; Score 39; DB 5; Length 1322;
Best Local Similarity 63.6%; Pred. No. 1e+02; 2; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LQVLGLOLPTP 12
DB 592 VOYIGLPAPTP 602

RESULT 9
ID O9NAL2 PRELIMINARY; PRT; 247 AA.
AC O9NAL2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Y113G7B.19 PROTEIN.
GN Y113G7B.19.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RP SEQUENCE FROM N.A.
RA Lennard N.;

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none:
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AL110477; CAB76742.1;
SO SEQUENCE 247 AA; 28826 MW; 199EA394CA1059B0 CRC64;

Query Match 63.3%; Score 38; DB 5; Length 247;
Best Local Similarity 70.0%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 ELQVGLQLPT 10
Db 222 KLOTGLQLPT 231

RESULT 10
Q9PB61 PRELIMINARY: PRT; 317 AA.
AC Q9PB61;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN XP2283.
GN XP2283.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xylella.
NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9ASC;
RX MEDLINE=20365717; Pubmed=10910347;
RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia J.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Coutinho L.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Facinanci A.P., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Fraga J.S., Franca S.C., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Villorelio C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Quagxo R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA de Souza A.P., Terenzi M.F., Truifi D., Tsai S.M., Tsubako M.H.,
RA Valida H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,
RL "The genome sequence of the plant pathogen Xylella fastidiosa."
DR EMBL: AF004040; AAF85082.1;
DR INTERPRO: IPR001279;
DR PFAM: PF00753; lactamase_B; 1.
KW Hypothetical protein.
SO SEQUENCE 317 AA; 34299 MW; A99CE5F8A6B557B9 CRC64;

Query Match 63.3%; Score 38; DB 2; Length 317;

Best Local Similarity 50.0%; Pred. No. 36;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 ELQVGLQLPT 12
Db 182 ELRLIGMDGPTP 193

RESULT 11
Q70399 PRELIMINARY: PRT; 501 AA.
AC Q70399;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE P2X2 RECEPTOR SPLICED VARIANT P2X2-3.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH STRAIN 2; TISSUE=ORGAN OF CORTI;
RA Parker M.S., Larroque M.L., Campbell J.M., Bobbin R.P., Deininger P.,
RL Hear. Res. 0:0-0(1998).
CC -1 SUBUNIT: HOMO- OR HETEROPOLYMERS (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC EMBL: AF053329; AAC08994.1;
DR EMBL: AF053329; AAC08994.1;
DR INTERPRO: IPR001429;
DR INTERPRO: IPR003045;
DR PFAM: PF00864; P2X2_receptor; 2.
DR PRINTS: PR01307; P2X2RECEPTOR.
DR PRINTS: PR01309; P2X2RECEPTOR.
DR PROSITE: PS01212; P2X2RECEPTOR; 1.
DR PRODOM: PD002383; -1.
KW Ionic channel; Transmembrane; Ion transport; Receptor.
SO SEQUENCE 501 AA; 55771 MW; 2B87EBB9F3CAC244 CRC64;

Query Match 63.3%; Score 38; DB 11; Length 501;
Best Local Similarity 58.3%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 ELQVGLQLPT 12
Db 450 QAKLGLQNPPT 461

RESULT 12
Q9NXB8 PRELIMINARY: PRT; 649 AA.
AC Q9NXB8;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE CDNA FLJ20335 FIS. CLONE HEP11429 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawakami T., Nozuchi S., Itoh T., Shigetani K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.,
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK000342; BAA91097.1;
DR NON_TER 649
FT 649
SO SEQUENCE 649 AA; 73060 MW; 18B9F87D373BC123 CRC64;

Query Match 63.3%; Score 38; DB 4; Length 649;
 Best Local Similarity 63.6%; Pred. No. 76;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LOVIGLOLPT 12
 ||:|||||
 Db 66 LOVIGLOLPT 76

RESULT 13
 044982

PRELIMINARY; PRT; 134 AA.

AC 044982;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE CSAG6.4 PROTEIN (FRAGMENT).
 GN CSAG6.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 ON NCBI_Taxid=6239;
 RX MEDLINE=94150718; PubMed=7906398;
 RC STRAIN-BRISTOL N2;
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
 Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 Jones M., Keshaw J., Kirsten J., Laister J., Latreille P.,
 Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 Parsons J., Percy C., Riklen L., Koopra A., Saunders S., Showkeen R.,
 Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RT Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Tin-Mollam A., Graves T., Ozerky P.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF043698; AAB97562.1;
 FT NOC_TER 1
 RT SEQUENCE 134 AA; 13720 MW; 24F44A10FA48CDEE CRC64;

Query Match 61.7%; Score 37; DB 5; Length 134;
 Best Local Similarity 70.0%; Pred. No. 22;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 QVLGLOLPT 12
 ||:|||||
 Db 98 QVLGLOLPT 107

RESULT 14
 084727

PRELIMINARY; PRT; 226 AA.

AC 084727;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PHOSPHOGLYCERATE MUTASE.
 GN PGM.
 OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 ON NCBI_Taxid=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UW-3/CX;
 RX MEDLINE=99000809; PubMed=9784136;
 RA Stephens R.S., Kalman S., Lammell C.J., Fan J., Marathe R., Aravind L.,
 Mitchell W.P., Olinger L., Tatunov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 Chlamydia trachomatis";
 RL Science 282:754-759(1998).
 DR EMBL: AF001343; AAC68317.1;
 DR HSP; P00950; 5PGM.
 DR INTERPRO: IPR001345;
 DR PFAM: PF00300; PGM; 2.
 DR PROSITE: PS00175; PG_MUTASE; UNKNOWN_1.
 SQ SEQUENCE 226 AA; 25785 MW; 82AE9C780D4008BD CRC64;

Query Match 61.7%; Score 37; DB 2; Length 226;
 Best Local Similarity 72.7%; Pred. No. 39;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ELOVIGLOLPT 11
 ||:|||||
 Db 195 ELOVIGLOLPT 205

RESULT 15
 09R201

PRELIMINARY; PRT; 253 AA.

AC 09R201;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE TRANSCRIPTIONAL REGULATOR, ICLR FAMILY.
 GN DRA0152.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 ON NCBI_Taxid=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,
 Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 Fraser C.M.;
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome Sequence of the Radioreistant Bacterium Deinococcus
 radiodurans R1";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001862; AAF1212.1;
 DR TIGR: DRA0152;
 DR INTERPRO: IPR000285;
 DR PFAM: PF01614; ICLR; 1.
 SQ SEQUENCE 253 AA; 27299 MW; DBB8F05ABE6F65A8 CRC64;

Query Match 61.7%; Score 37; DB 2; Length 253;
 Best Local Similarity 63.6%; Pred. No. 44;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LOVIGLOLPT 12
 ||:|||||
 Db 205 LALGVSLPT 215

Search completed: June 4, 2001, 12:23:31

Tue Jun 5 07:09:45 2001

Job time: 594 sec

us-09-284-107-28.rspt

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:13:33 ; Search time 106.14 Seconds
(without alignments)
6.463 Million cell updates/sec

Title: US-09-284-107-29

Sequence: 1 VWFHVLFLAVG 12

Scoring table: BIOSUM62
Gapop 10.0, Capext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_0401.*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT.*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT.*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT.*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT.*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT.*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT.*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT.*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT.*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT.*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT.*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT.*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT.*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT.*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT.*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT.*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT.*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT.*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT.*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT.*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT.*
22: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT.*
23: /SIDS6/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	12	19 W60561	Oligopeptide from
2	68	100.0	344	17 R91439	Human FCRI (CDNA C
3	68	100.0	344	21 Y96183	Human macrophage-s
4	68	100.0	374	17 R91438	Human FCRI (CDNA C
5	68	100.0	374	17 W00859	Human FCRI (CDNA C
6	68	100.0	374	19 W80448	Human FC receptor
7	68	100.0	374	19 W97833	Human FC receptor
8	68	100.0	374	19 W97834	Human FC receptor
9	68	100.0	374	21 Y96134	Human macrophage-s
10	68	100.0	374	21 Y96226	Human high affinity
11	68	100.0	399	21 B43683	Human cancer assoc

12	64	94.1	410	12 R12428	Hybrid FC(gamma)RI
13	51	75.0	377	20 W86195	Human FC receptor
14	49	72.1	243	21 G50156	Arabidopsis thalia
15	49	72.1	257	21 G50155	Arabidopsis thalia
16	49	72.1	373	21 G50154	Arabidopsis thalia
17	42	61.8	96	21 G40831	Zea mays protein f
18	42	61.8	98	21 G40830	Zea mays protein f
19	42	61.8	164	21 G40829	Zea mays protein f
20	38.5	56.6	1084	21 B42371	Human ORFX ORF2135
21	38	55.9	65	21 G37550	Arabidopsis thalia
22	38	55.9	203	21 G37550	Arabidopsis thalia
23	38	55.9	266	20 Y16106	A formate transpor
24	38	55.9	266	21 G06465	Arabidopsis thalia
25	38	55.9	266	21 G06465	Arabidopsis thalia
26	38	55.9	266	21 G37549	Arabidopsis thalia
27	38	55.9	267	21 G05036	Arabidopsis thalia
28	38	55.9	267	21 G14200	Arabidopsis thalia
29	38	55.9	267	21 G50429	Arabidopsis thalia
30	38	55.9	273	21 G06464	Arabidopsis thalia
31	38	55.9	273	21 G37548	Arabidopsis thalia
32	38	55.9	281	21 G14199	Arabidopsis thalia
33	38	55.9	327	21 G05035	Arabidopsis thalia
34	38	55.9	327	21 G50428	Arabidopsis thalia
35	38	55.9	391	21 G42852	Arabidopsis thalia
36	38	55.9	391	21 G42852	Arabidopsis thalia
37	38	55.9	395	20 Y24478	Nicotiana panicula
38	38	55.9	398	20 Y24479	Nicotiana panicula
39	38	55.9	398	21 G05034	Arabidopsis thalia
40	38	55.9	398	21 G50427	Arabidopsis thalia
41	38	55.9	423	21 G09391	Arabidopsis thalia
42	38	55.9	423	21 G42851	Arabidopsis thalia
43	38	55.9	610	19 W38410	Yeast acyl-coenzyme
44	38	55.9	610	19 W38417	Yeast acyl-coenzyme
45	38	55.9	1176	21 G49881	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
ID	W60561
W60561	standard; peptide: 12 AA.
AC	W60561:
XX	18-AUG-1998 (first entry)
XX	Oligopeptide from extracellular domain of CD64.
XX	Extracellular domain; CD64; identification; antibody;
KW	Immunohistochemical; immunofluorescent analysis; detection;
KW	cell transformation; mutation; anti; oncogene.
XX	Synthetic.
OS	W09815833-A1.
XX	16-APR-1998.
XX	07-OCT-1997; 97WO-NL00557.
XX	08-OCT-1996; 96EP-0202791.
XX	(UYUT-) RIJUSUNIV UTRECHT.
XX	De Krul CA, Logtenberg T;
XX	WPI: 1998-240964/21.
XX	Identifying peptide(s) binding specifically to protein target - by
XX	expressing on phage surface and testing for binding to immobilised
XX	oligo:peptide derived from the target. useful for, e.g. Identifying
XX	specific antibodies

XX Example 1: Page 29; 40pp; English.

PS Synthetic oligopeptides W60537-61 are derived from the extracellular
CC domain of C664. They were synthesized on 25 polyethylene rods as
CC adjacent, non-overlapping sets. Peptide specific phages were isolated by
CC affinity binding to the C664-covered rods. The specification describes a
CC method for the identification of a peptide able to bind specifically to a
CC target protein. The method comprises displaying the peptide on the
CC surface of a replicable display package, synthesizing oligopeptides
CC derived from the target protein on a solid phase, and testing for binding
CC between the peptide and oligopeptides. The method is used to screen large
CC peptide libraries, especially to detect antibodies, or their fragments,
CC that bind to cell markers or that can differentiate between different
CC forms of the same protein, including bispecific antibodies that bind to
CC two non-overlapping epitopes on the same monomeric antigen or two
CC epitopes on different molecules. The genes/oligonucleotides that encode
CC selected peptides can be isolated and used for recombinant production of
CC antibodies. Selected antibodies are used, e.g. for immunohistochemical or
CC immunofluorescent analysis, and also to detect cell transformation caused
CC by mutation in (anti)oncogenes.

XX SQ Sequence 12 AA:

Query Match 100.0%; Score 68; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 1,1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHVLFTYLAAG 12
DB 1 vwfhlvlyfayg 12

RESULT 2
R91439 R91439 standard; Protein; 344 AA.

AC R91439;

DT 30-OCT-1996 (first entry)

DE Human FCRI (cDNA clone p98 product).

KW Cell surface antigen; cloning; immunoselection; immunotherapy;
therapy; diagnosis; vector; FCRI; Fc receptor.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 58 /note= "amino acid 58 is Leu in clone p135
FT translated product"

PN US5506126-A.

PD 09-APR-1996.

PF 25-FEB-1988; 88US-0160416.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 18-OCT-1993; 93US-0139273.

PA (GEHO) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

DR WPI; 1996-200279/20.

XX N-PSDB; T14718.

PT Cloning of cDNA encoding cell surface antigen - useful for isolation
PT of diagnostic and therapeutic proteins

XX Example 10: Column 55-56; 79pp; English.

CC The amino acid sequence (R91439) of human FCRI was detd. from a
CC cDNA clone, p98 (T14718), obtd. from a cDNA library using an
CC immunoselection cloning method. FCRI is a high affinity receptor
CC for the Fc portion of IgG, normally located on cell surfaces of
CC macrophages. 2 other cDNA clones, p135 (T14717) and p90 (T14719),
CC coded for a variants of the FCRI sequence; the C-terminal sequence
CC of the p98 product is truncated compared with those of the p135
CC and p90 products (see also R91438 and W008529). A fusion protein
CC of FCRI and a receptor ligand will be helpful in increasing the
CC potency of antibodies in therapy.

XX SQ Sequence 344 AA:

Query Match 100.0%; Score 68; DB 17; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHVLFTYLAAG 12
DB 289 vwfhlvlyfayg 300

RESULT 3
Y96183 Y96183 standard; Protein; 344 AA.

AC Y96183;

DT 19-DEC-2000 (first entry)

DE Human macrophage-specific FCRI.

KW Macrophage; FCRI; cell surface antigen; human; immunoselection;
panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;

KW immune disorder; infection; asthma; immune-complex disease;

XX amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 15 /note= "encoded by GCG"

FT Misc-difference 38 /note= "encoded by ACC"

FT Misc-difference 50 /note= "encoded by CCG"

FT Misc-difference 51 /note= "encoded by ACC"

FT Misc-difference 55 /note= "encoded by ACC"

FT Misc-difference 56 /note= "encoded by CAC"

FT Misc-difference 60 /note= "encoded by TCC"

FT Misc-difference 64 /note= "encoded by CCC"

FT Misc-difference 82 /note= "encoded by CAG"

FT Misc-difference 116 /note= "encoded by CAA"

FT Misc-difference 117 /note= "encoded by GAA"

FT Misc-difference 122 /note= "encoded by GCA"

FT Misc-difference 123 /note= "encoded by TTC"

FT Misc-difference 126 /note= "encoded by ACC"

```

CC protein of FCRI and a receptor ligand will be helpful to increase
CC the potencies of antibodies in therapy.
XX
S0 Sequence 344 AA;

Query Match 100.0%; Score 68; DB 21; Length 344;
Best Local Similarity 100.0%; Pred. NO. 0.00038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VWFHVLFLYLAVG 12
   |||||
DB 289 vwfhlvflayvg 300

RESULT 4
R91438
XX ID R91438 standard; protein; 374 AA.
XX AC R91438;
XX
XX 30-OCT-1996 (first entry)
XX
XX Human FCRI (cDNA clone p135 product).
XX
XX Cell surface antigen; cloning; immunoselection; immunotherapy;
XX KW therapy; diagnosis; vector; FCRI; Fc receptor.
XX RM
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 25 /note= "amino acid 25 is Thr in clone p90 and p98
FT FT translated products"
FT FT
FT FT Misc-difference 58 /note= "amino acid 58 is Val in p90 clone
FT FT translated product"
XX
XX US5506126-A.
XX
XX 09-APR-1996.
XX
XX 25-FEB-1988; 88US-0160416.
XX
XX 01-DEC-1992; 92US-0983647.
XX PR 25-FEB-1988; 88US-0160416.
XX PR 13-JUL-1989; 89US-0379076.
XX PR 13-JUL-1990; 90US-0553759.
XX PR 18-OCT-1993; 93US-0139273.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Aruffo A, Seed B;
XX
XX WPI: 1996-200279/20.
XX DR N-PSDB; T14717.
XX
XX Cloning of cDNA encoding cell surface antigen - useful for isolation
XX of diagnostic and therapeutic proteins
XX
XX Example 10; Column 55-56; 79pp; English.
XX
XX The amino acid sequence (R91438) of human FCRI was detd. from a
XX cDNA clone, p135 (T14717), obtd. from a cDNA library using an
XX immunoselection cloning method. FCRI is a high affinity receptor
XX for the Fe portion of IgG, normally located on cell surfaces of
XX macrophages. Another isolated cDNA clone, p90 (T14719), coded
XX for a variant (W00859) of the FCRI sequence, and a third clone, p98
XX (T14718), coded for an FCRI (R91439) having a different C-terminal
XX sequence. A fusion protein of FCRI and a receptor ligand will
XX be helpful in increasing the potency of antibodies in therapy.
XX
S0 Sequence 374 AA;

```

Query Match 100.0%; Score 68; DB 17; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.00042; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WVEHVLFLAVG 12
|||||
DB 289 wvhvlflavg 300

RESULT 5

W00859 W00859 standard; Protein; 374 AA.

AC W00859;

DT 30-OCT-1996 (first entry)

DE Human FCRI (cDNA clone p90 product).

KW Cell surface antigen; cloning; immunoselection; immunotherapy;

therapy; diagnosis; vector; FCRI; Fc receptor.

OS Homo sapiens.

PN US5506126-A.

PD 09-APR-1996.

FE 25-FEB-1988; 88US-0160416.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 18-OCT-1993; 93US-0139273.

PA (GCHO) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

DR WPI: 1996-200279/20.

DR N-PSDB: T14719.

PT Cloning of cDNA encoding cell surface antigen - useful for isolation

PT of diagnostic and therapeutic proteins

PS Example 10; Column 55-56; 79pp; English.

XX The amino acid sequence (W00859) of human FCRI was detd. from a

CC cDNA clone, p90 (T14719), obtd. from a cDNA library using an

CC immunoselection cloning method. FCRI is a high affinity receptor

CC for the Fc portion of IgG, normally located on cell surfaces of

CC macrophages. Other cDNA clones (see also T14717-18) coded for

CC variants (R01438-39) of the FCRI sequence. A fusion protein of

CC FCRI and a receptor ligand will be helpful in increasing the

CC potency of antibodies in therapy.

CC Sequence 374 AA:

ID W80448 standard; Protein; 374 AA.

XX W80448;

AC 07-JUN-1999 (first entry)

DE Human Fc receptor I.

KW Fc receptor I; FCRI; cell surface antigen; lymphocyte; human;

KW cloning.

OS Homo sapiens.

PN Key Location/Qualifiers

FT Misc-difference 2 /note= "encoded by TGC"

FT Misc-difference 23 /note= "encoded by CTC"

FT Misc-difference 44 /note= "encoded by GAG"

FT Misc-difference 45 /note= "encoded by CTG"

FT Misc-difference 60 /note= "encoded by CCC"

FT Misc-difference 77 /note= "encoded by AAT"

FT Misc-difference 85 /note= "encoded by TCC"

FT Misc-difference 99 /note= "encoded by CAA"

FT Misc-difference 103 /note= "encoded by CCC"

FT Misc-difference 141 /note= "encoded by GGC"

FT Misc-difference 159 /note= "encoded by AAC"

FT Misc-difference 171 /note= "encoded by ATG"

FT Misc-difference 176 /note= "encoded by GTC"

FT Misc-difference 236 /note= "encoded by GCG"

PN US5830731-A.

PD 03-NOV-1998.

FE 21-MAY-1997; 97US-0861205.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 21-MAY-1997; 97US-0861205.

PA (GCHO) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

DR WPI: 1998-609251/51.

DR N-PSDB: V63456.

PT New cloning vector and polylinker - based on existing sequences for

PT efficient cloning and expression of mammalian cDNA(s), especially

PT human lymphocyte antigenic sequences

PS Example 10; Column 53-54; 75pp; English.

XX This is the amino acid sequence of human Fc receptor I (FCRI), as

CC deduced from cDNA clone p135 (see V63456) isolated using a rapid

CC immunoselection cloning method from a cDNA library expressed in COS

CC cells. The cDNA library was constructed from polyA RNA of cells

CC from a single patient undergoing extracorporeal interleukin-2

CC induction therapy. Clones p90 (see X07372) and p98 (see X07373)
 CC were also obtained. A novel method for cloning cDNAs from mammalian
 CC expression libraries is based on transient expression of an antigen
 CC in eukaryotic cells and selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. The method is useful for
 CC the isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell. It
 CC has been used to clone genes (see V63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see W80440-55). The isolated
 CC genes can be expressed in a prokaryotic or eukaryotic host cells to
 CC produce the encoded protein. The invention also provides high
 CC efficiency expression vectors (see V63441 and V63444) which allow
 CC the generation of very large mammalian expression libraries. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.

SO Sequence 374 AA;

Query Match 100.0%; Score 68; DB 19; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.00042; Mismatches 0; Indels 0; Gaps 0;

Matches 12; Conservative 0;

Oy 1 W97833 standard; Protein; 374 AA.

Db 289 wfhvlfy1av9 300

RESULT 7

W97833

W97833; 07-JUN-1999 (first entry)

DE Human Fc receptor I.

FC receptor I; FCRI; cell surface antigen; lymphocyte; human;

cloning.

Homo sapiens.

Location/Qualifiers

Misc-difference 2

Misc-difference 23

Misc-difference 44

Misc-difference 45

Misc-difference 60

Misc-difference 77

Misc-difference 85

Misc-difference 99

Misc-difference 103

Misc-difference 141

Misc-difference 159

Misc-difference 171

Misc-difference 176

Misc-difference 256

/note- "encoded by GGC"

US5830731-A.

03-NOV-1998.

21-MAY-1997; 97US-0861205.

01-DEC-1992; 92US-0983647.

25-FEB-1988; 88US-0160416.

13-JUL-1989; 89US-0379076.

13-JUL-1990; 90US-0553759.

21-MAY-1997; 97US-0861205.

(GENO) GEN HOSPITAL CORP.

Aruffo A, Seed B;

WPI: 1998-609251/51.

N-PSDB: X07372.

New cloning vector and polylinker - based on existing sequences for

efficient cloning and expression of mammalian cDNA(s), especially

human lymphocyte antigenic sequences

Example 10; Column 53-54; 75pp; English.

This is the amino acid sequence of human Fc receptor I (FCRI), as

deduced from cDNA clone p90 (see X07372) isolated using a rapid

immunoscreening cloning method from a cDNA library expressed in COS

cells. The cDNA library was constructed from polyA RNA of cells

from a single patient undergoing extracorporeal interleukin-2

induction therapy. Clones p135 (see V63456) and p98 (see X07373)

were also obtained. A novel method for cloning cDNAs from mammalian

expression libraries is based on transient expression of an antigen

in eukaryotic cells and selection of cells expressing the antigen by

adhesion to an antibody-coated substrate. The method is useful for

the isolation and cloning of any protein which can be expressed and

transported to the cell surface membrane of a eukaryotic cell. It

has been used to clone genes (see V63442-63) encoding cell surface

antigens from mammalian lymphocytes (see W80440-55). The isolated

genes can be expressed in a prokaryotic or eukaryotic host cells to

produce the encoded protein. The invention also provides high

efficiency expression vectors (see V63441 and V63444) which allow

the generation of very large mammalian expression libraries. The

purified genes and proteins are useful for immunodiagnostic and

immunotherapeutic applications, including the diagnosis and

treatment of immune-mediated infections, diseases, and disorders of

animals, including humans.

Sequence 374 AA;

Query Match 100.0%; Score 68; DB 19; Length 374;

Best Local Similarity 100.0%; Pred. No. 0.00042;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 W97833 standard; Protein; 374 AA.

Db 289 wfhvlfy1av9 300

RESULT 8

W97833

W97833; 07-JUN-1999 (first entry)

DE Human Fc receptor I.

FC receptor I; FCRI; cell surface antigen; lymphocyte; human;

Query Match	100.0%	Score 68	DB 19	Length 374
Best Local Similarity	100.0%	Pred. No. 0.00042		
Matches 12	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1 VWFHVLFIYLAAG 12			
DB	289 VWFHVLFIYLAAG 300			
RESULT	9			
ID	Y96134			
XX	Y96134 standard; Protein; 374 AA.			
AC	Y96134;			
XX				
DT	19-DEC-2000 (first entry)			
XX				
DE	Human macrophage-specific FCRI.			
KW	Macrophage; FCRI; cell surface antigen; human; immunoselection;			
KW	painning; immunodiagnosis; diagnosis; immunotherapy; gene therapy;			
KW	immune disorder; infection; asthma; immune-complex disease;			
XX	amyloidosis; multiple sclerosis; parasitic disease; Fc receptor.			
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
FH	Misc-difference 15			
FT	/note= "encoded by GCG"			
FT	Misc-difference 38			
FT	/note= "encoded by ACC"			
FT	Misc-difference 50			
FT	/note= "encoded by CGG"			
FT	Misc-difference 51			
FT	/note= "encoded by ACC"			
FT	Misc-difference 55			
FT	/note= "encoded by CAC"			
FT	Misc-difference 56			
FT	/note= "encoded by TTC"			
FT	Misc-difference 60			
FT	/note= "encoded by CCC"			
FT	Misc-difference 64			
FT	/note= "encoded by CAG"			
FT	Misc-difference 82			
FT	/note= "encoded by CAA"			
FT	Misc-difference 116			
FT	/note= "encoded by GAA"			
FT	Misc-difference 117			
FT	/note= "encoded by GCA"			
FT	Misc-difference 122			
FT	/note= "encoded by TTC"			
FT	Misc-difference 123			
FT	/note= "encoded by ACC"			
FT	Misc-difference 126			
FT	/note= "encoded by CCG"			
FT	Misc-difference 129			
FT	/note= "encoded by CAT"			
FT	Misc-difference 134			
FT	/note= "encoded by AAT"			
FT	Misc-difference 136			
FT	/note= "encoded by GTT"			

FT Misc-difference 139 /note= "encoded by CCA"
FT Misc-difference 140 /note= "encoded by AAT"
FT Misc-difference 213 /note= "encoded by CAA"
FT Misc-difference 216 /note= "encoded by TTC"
FT Misc-difference 220 /note= "encoded by CGT"
FT Misc-difference 268 /note= "encoded by AAT"
FT Misc-difference 305 /note= "encoded by GTG"
FT Misc-difference 306 /note= "encoded by AAC"
FT Misc-difference 332 /note= "encoded by GGT"
FT US6111093-A.
PN
XX
PD 29-AUG-2000.
XX
PF 28-OCT-1998; 98US-0181612..
XX
PR 01-DEC-1992; 92US-0983647.
PR 25-FEB-1988; 88US-0160416.
PR 13-JUL-1989; 89US-0379076.
PR 23-MAR-1990; 90US-0498809.
PR 13-JUL-1990; 90US-0553759.
XX
XA (GEHO) GEN HOSPITAL CORP.
PA
PI Stamenkovic I, Seed B;
PT WPI: 2000-586382/55.
DR N-PSDB; A50592.
XX
XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
PT useful for immunodiagnosis and immunotherapy of immune-mediated
PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
PT diseases
XX
XX Example 10; Column 53-55; 75pp; English.

The present sequence is that of a human macrophage specific FCRI, as deduced from cDNA clone p135 (see A50592), which was isolated from a cDNA library expressed in COS cells using a novel method of the invention designed to isolate CSA nucleic acids. The method is based upon transient expression of a CSA in eukaryotic cells and physical selection of cells expressing the antigen by adhesion to (panning on) an antibody-coated substrate such as a culture dish. CSA nucleic acids isolated by the method of the invention, and the proteins they encode, are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders in animals, including humans. These disorders include asthma, immune-complex disease, amyloidosis, parasitic diseases or multiple sclerosis. FCRI is a high affinity receptor for the Fc portion of IgG, normally located on the cell surfaces of macrophages. The ability to interfere with such bonding, or to cause it to occur on surfaces other than macrophages, is useful in therapy. A fusion protein of FCRI and a receptor ligand will be helpful to increase the potencies of antibodies in therapy.

Db	289	wfhwlfyfav	300
RESULT	10		
ID	Y96226		
AC	Y96226 standard; Protein; 374 AA.		
XX	Y96226;		
DT	11-SEP-2000 (first entry)		
DE	Human high affinity Fc receptor, Fcgammar1.		
XX			
XX	Human: high affinity Fc receptor; Fcgammar1; immunoglobulin;		
KM	infection; immune response; CD64; monocyte; macrophage; neutrophil;		
KM	eosinophil; HIV; IgG; immunosuppressive; antineutrotic; cytostatic;		
KM	antihistaminic; AIDS; SLE; multiple myeloma; rheumatoid arthritis;		
KM	systemic lupus erythematosus; tumour.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1006183-A1.		
XX			
PD	07-JUN-2000.		
XX			
PF	03-DEC-1998; 98EP-0122969.		
XX			
PR	03-DEC-1998; 98EP-0122969.		
XX			
PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.		
XX			
DR	WPI: 2000-367966/32.		
DR	N-PSDB; A27466.		
PT	Novel Fc receptor lacking transmembrane domains, a signal peptide, and		
PT	glycosylation, useful for diagnosing and treating immune disorders and		
XX	cancer		
XX			
PS	Disclosure; Page 26-28; 60pp; English.		
XX			
XX	The present sequence is the human high affinity Fc receptor, Fcgammar1.		
CC	Fcgammar1 is also known as CD64. Fc receptors play an important		
CC	role in defending the body against infections. First, pathogens are		
CC	opsonised by serum immunoglobulins. The resulting complex then binds to		
CC	cells expressing Fc receptors. Fcgammar1 molecules are expressed by		
CC	monocytes and macrophages, but expression can also be induced on		
CC	neutrophils and eosinophils. Upon Fc receptor activation, immune effector		
CC	pathways are activated, leading to immune response. The present sequence		
CC	may be modified to produce recombinant versions. The recombinant Fc		
CC	receptor consist only of the extracellular portion of the receptor and		
CC	are not glycosylated i.e. they do not have transmembrane domains or		
CC	signal peptides. The recombinant proteins may be used in immunoassays to		
CC	determine the immune status of patients with chronic diseases of the		
CC	immune system, e.g. AIDS, systemic lupus erythematosus (SLE), multiple		
CC	myeloma (MM), or rheumatoid arthritis. In addition, pharmaceutical		
CC	compositions containing recombinant proteins may be used to treat or		
CC	prevent autoimmune diseases, allergies or tumours, especially AIDS,		
CC	rheumatoid arthritis or MM.		
XX			
XX	Sequence 374 AA;		
XX			
Query Match	100.0%;	Score 68;	DB 21; Length 374;
Best Local Similarity	100.0%;	Pred. NO. 0.00042;	
Matches 12; Conservative	0;	Mismatches	0; Indels
		Gaps	0;

Query Match	100.0%;	Score 68;	DB 21;	Length 374;
Best Local Similarity	100.0%;	Pred. No. 0.00042;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
0y	1 VMPHVLFLYAVG 12			

QY	1	WVFHVLFLYLAvg	12
Db	289	wfHVLFLYLAvg	300
RESULT	11		

B43683
ID B43683 standard; Protein: 399 AA.
AC B43683;
XX
XX
DT 08-FEB-2001 (first entry)
DE Human cancer associated protein sequence SEQ ID NO:1128.
XX
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
antidiabetic; antihemetic; antihemetic; antihemetic; antiviral;
antitumor; antitumor; antitumor; antitumor; antitumor;
dermatologic; neuroprotective; thrombolytic; coagulant; nocitopic;
vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;
immune disorder; haematopoietic cell disorder; autoimmune disorder;
allergic reaction; graft versus host disease; organ rejection;
haemostatic; thrombolytic; cardiovascular disorder; infection;
neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN MO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587533/55.
DR N-PSDB: C77892.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 11; Page 1739-1740; 2352pp; English.
XX
CC C77607 to C78448 encode the human cancer associated proteins given in
CC B43398 to B44239. The proteins can have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC cytoskeletal; proliferative; vulnery; immunomodulator; antidiabetic;
CC antihemetic; antihemetic; antihemetic; antihemetic;
CC antitumor; antitumor; antitumor; antitumor;
CC neuroprotective; cardiant; thrombolytic; coagulant; nocitopic;
CC vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;
CC polypeptides can be used for preventing, treating or ameliorating medical
CC conditions and diagnosing pathological conditions. Polynucleotides
CC polypeptides, antibodies, agonists and antagonists from the present
CC invention may be used to treat immune disorders by activating or
CC inhibiting the proliferation, differentiation or mobilisation of immune
CC cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC allergic reactions, graft versus host disease and organ rejection.
CC modulate haemostatic or thrombolytic activity, modulate inflammation,
CC cancers, cardiovascular disorders, neurological disease and bacterial or
CC viral infections. The peptides, nucleotides, antibodies, agonists and
CC antagonists may be also be used in drug screens. C78449 to C78457 and
CC B44240 represent sequences used in the exemplification of the present
CC invention.
XX
SQ - Sequence 399 AA;

Query Match 100.0%; Score 68; DB 21; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHVLFLAVG 12
|||||

DB 314 vwfhlflavg 325
RESULT 12
R12428
ID R12428 standard; Protein: 410 AA.
XX
XX
AC R12428;
XX
XX
DT 14-AUG-1991 (first entry)
DE Hybrid Fc(gamma)RI/II receptor.
XX
XX
KW Immunoglobulin-binding molecule; Fc(gamma)RI; Fc(gamma)RII.
XX
FH Key Location/Qualifiers
FT Region 1..203
FT /label= Fc(gamma)RI domains d1 and d2
FT 204..410
FT /label= Fc(gamma)RI D3, transmembrane and
FT cytoplasmic domains
XX
XX
PN MO9106570-A.
XX
PD 16-MAY-1991.
XX
PF 25-OCT-1990; 90WO-A000513.
XX
PR 25-OCT-1989; 89AU-0007045.
XX
PA (UYME-) UNIV MELBOURNE.
XX
PI Hogarth PM, Hulett MD, Ierino FL, McKenzie IF, Osman N;
XX
DR WPI: 1991-164135/22.
DR N-PSDB: Q12079.
XX
PT New immunoglobulin binding hybrid Fc receptor molecules - used to
PT control auto-immune diseases and allergic reactions and to regulate
PT antibody prodn.
XX
PS Claim 10; fig 3B; 80pp; English.
XX
CC This chimeric sequence is one example of a hybrid Fc receptor
CC of the invention. The hybrid Fc receptor retains the Ig binding
CC functions of the constituent receptor types. It can be used for
CC regulating antibody production in vivo and in plasmaphoresis to
CC remove immune complexes or pathological antibodies. Rabbit IgG
CC bound to cells transfected with the chimeric Fc(gamma) receptor.
CC See also Q12078.
XX
SQ Sequence 410 AA;

Query Match 94.1%; Score 64; DB 12; Length 410;
Best Local Similarity 83.3%; Pred. No. 0.0022;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHVLFLAVG 12
|||||

DB 304 vwfhlflavsg 315

RESULT 13
W86195
ID W86195 standard; Protein: 377 AA.
XX
XX
AC W86195;
XX
DT 10-MAY-1999 (first entry)
XX
XX
DE Human Fc receptor I.

XX	Fc receptor I; CD36; cell surface antigen; human; cDNA library.
XW	
OS	Homo sapiens.
FX	
FT	Key
FT	Protein
FT	Location/Qualifiers
FT	1..377
FT	/note= "this sequence contains a considerable number of differences from the sequence deduced from the DNA sequence given in the specification"
FT	
PN	US5849898-A.
PD	
PD	15-DEC-1998.
XP	
XP	07-JUN-1995;
PE	95US-0485447.
XX	
PR	01-DEC-1992;
PR	92US-0983647.
PR	25-FEB-1988;
PR	88US-0160416.
PR	13-JUL-1989;
PR	89US-0379076.
PR	23-MAR-1990;
PR	90US-0498809.
PR	13-JUL-1990;
PR	90US-0553759.
PR	07-JUN-1995;
PR	95US-0485447.
PA	(GENO) GEN HOSPITAL CORP.
PI	Allen J, Amiot M, Aruffo A, Cemerini D, Laufer L,
PI	Oquendo C, Seed B, Simmons D, Stamenkovic I, Stengelin S;
XX	
DR	WPI: 1999-069813/06./
DR	N-PSSB: V81213.
PT	cDNA encoding human CD40 antigen - useful for cloning cDNA encoding cell surface antigens, constructing cDNA libraries, expression vectors for expression in eukaryotic cells or their fragments
PS	Example 10; Column 55-56; 79pp: English.
XX	
CC	This polypeptide human Fc receptor I (FCRI). FCRI cDNAs (see V81213) were isolated using a rapid immunoselection cloning method from a cDNA library expressed in COS cells. The cDNA library was constructed from polyA RNA of cells from a single patient undergoing extracorporeal interleukin-2 induction therapy. DNA sequence analysis revealed that the cDNAs encoded type I integral membrane proteins with 3 extracellular immunoglobulin domains. The invention provides a novel method for cloning cDNAs from mammalian expression libraries. This is based on transient expression of an antigen in eukaryotic cells and physical selection of cells expressing the antigen by adhesion to an antibody-coated substrate. The method is useful for the isolation and molecular cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell. CD40 cDNA (see V81198) is specifically claimed.
SO	Sequence 377 AA:
QY	
Query Match	75.0%; Score 51; DB 20; Length 377;
Best Local Similarity	90.0%; Pred. No. 0.36;
Matches 9; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
DB	
292 wfvnlfy1a 301	
I VFHFVLYIA 10	
RESULT 14	
G50156	
ID	G50156 standard; Protein; 243 AA.
AC	
G50156;	
BT	18-OCT-2000 (first entry)

XX	Arabidopsis thaliana protein fragment SEQ ID NO: 63527.
DE	Arabidopsis thaliana protein identification; signal transduction pathway; metabolic pathway;
XX	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
PD	
XX	06-SEP-2000.
PF	
XX	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	16-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139452.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139460.
PR	18-JUN-1999; 99US-0139461.
PR	18-JUN-1999; 99US-0139462.
PR	18-JUN-1999; 99US-0139463.
PR	18-JUN-1999; 99US-0139750.
PR	18-JUN-1999; 99US-0139763.

```

PR 21-JUN-1999; 99US-0139817
PR 22-JUN-1999; 99US-0139899
PR 23-JUN-1999; 99US-0140353
PR 23-JUN-1999; 99US-0140354
PR 24-JUN-1999; 99US-0140695
PR 28-JUN-1999; 99US-0140823
PR 29-JUN-1999; 99US-0140991
PR 30-JUN-1999; 99US-0141287
PR 01-JUL-1999; 99US-0141842
PR 02-JUL-1999; 99US-0142154
PR 06-JUL-1999; 99US-0142390
PR 08-JUL-1999; 99US-0142803
PR 09-JUL-1999; 99US-0142920
PR 12-JUL-1999; 99US-0142927
PR 13-JUL-1999; 99US-0143542
PR 14-JUL-1999; 99US-0143624
PR 15-JUL-1999; 99US-0144005
PR 16-JUL-1999; 99US-0144086
PR 19-JUL-1999; 99US-0144325
PR 19-JUL-1999; 99US-0144331
PR 19-JUL-1999; 99US-0144332
PR 19-JUL-1999; 99US-0144333
PR 19-JUL-1999; 99US-0144334
PR 19-JUL-1999; 99US-0144335
PR 20-JUL-1999; 99US-0144352
PR 20-JUL-1999; 99US-0144632
PR 20-JUL-1999; 99US-0144884
PR 21-JUL-1999; 99US-0144814
PR 21-JUL-1999; 99US-0145086
PR 21-JUL-1999; 99US-0145088
PR 22-JUL-1999; 99US-0145085
PR 22-JUL-1999; 99US-0145087
PR 22-JUL-1999; 99US-0145089
PR 22-JUL-1999; 99US-0145182
PR 23-JUL-1999; 99US-0145145
PR 23-JUL-1999; 99US-0145218
PR 23-JUL-1999; 99US-0145224
PR 26-JUL-1999; 99US-0145276
PR 27-JUL-1999; 99US-0145913
PR 27-JUL-1999; 99US-0145918
PR 27-JUL-1999; 99US-0145919
PR 28-JUL-1999; 99US-0145951
PR 02-AUG-1999; 99US-0146386
PR 02-AUG-1999; 99US-0146388
PR 03-AUG-1999; 99US-0146389
PR 04-AUG-1999; 99US-0147038
PR 04-AUG-1999; 99US-0147204
PR 04-AUG-1999; 99US-0147302
PR 05-AUG-1999; 99US-0147192
PR 06-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147303
PR 09-AUG-1999; 99US-0147416
PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935
PR 10-AUG-1999; 99US-0148171
PR 11-AUG-1999; 99US-0148319
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148565
PR 13-AUG-1999; 99US-0148684
PR 16-AUG-1999; 99US-0149368
PR 17-AUG-1999; 99US-0149368
PR 18-AUG-1999; 99US-0149175
PR 20-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149722
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149929
PR 23-AUG-1999; 99US-0149902
PR 23-AUG-1999; 99US-0149930
PR 25-AUG-1999; 99US-0150566
PR 26-AUG-1999; 99US-0150884
PR 27-AUG-1999; 99US-0151065
PR 27-AUG-1999; 99US-0151066

```

```

PR 27-AUG-1999; 99US-0151080
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151438
PR 01-SEP-1999; 99US-0151930
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153070
PR 13-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157753
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0158029
PR 08-OCT-1999; 99US-0158232
PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0159293
PR 13-OCT-1999; 99US-0159294
PR 13-OCT-1999; 99US-0159295
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159884
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160989
PR 23-OCT-1999; 99US-0161404
PR 25-OCT-1999; 99US-0161405
PR 25-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161359
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161993
PR 29-OCT-1999; 99US-0162142

```

```

Query Match 72.1% Score 49; DB 21; Length 243;
Best Local Similarity 54.5%; Pred. No. 0.49;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 WHEHVEFLAV 11
Db 219 iWfhllyeffl 229

```

```

RESULT 15
ID G50155 standard; Protein; 257 AA.
AC G50155;

```

```

DT 18-OCT-2000 (first entry)

```

```

DE Arabidopsis thaliana protein fragment SEQ ID NO: 63326.

```

```

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX

```

OS Arabidopsis thaliana.
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132408.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 72.1%; Score 49; DB 21; Length 257;
 Best Local Similarity 54.5%; Pred. No. 0.53;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VWFHLYFIAY 11
 :||:|:|:|:
 Db 233 lwfhlyfiffi 243

Search completed: June 4, 2001, 12:13:33
 Job time: 206 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:14:33 ; Search time 58.74 Seconds
(without alignments)
3.925 Million cell updates/sec

Title: US-09-284-107-29

Sequence: 1 VWFHVFYLAAG 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCYUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfilest.pep: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35.5	52.2	453	1	US-08-439-131A-5
2	35.5	52.2	453	1	US-08-440-674-4
3	35	51.5	361	1	US-08-415-751-36
4	34	50.0	162	1	US-08-624-125-5
5	34	50.0	162	1	US-08-624-125-8
6	34	50.0	162	2	US-08-663-310-4
7	34	50.0	162	2	US-08-663-310-11
8	34	50.0	162	2	US-09-006-491-4
9	34	50.0	162	2	US-09-006-491-11
10	34	50.0	162	4	US-09-335-919-4
11	34	50.0	162	4	US-09-335-919-11
12	34	50.0	261	2	US-08-332-562A-133
13	34	50.0	340	1	US-08-118-270-49
14	34	50.0	340	5	PCT-US93-08528-49
15	34	50.0	372	4	US-07-937-609-20
16	34	50.0	372	4	US-08-029-170-20
17	34	50.0	407	2	US-08-390-000A-6
18	34	50.0	407	5	PCT-US92-06532-7
19	34	50.0	667	4	US-08-959-004-11
20	33	48.5	234	1	US-08-287-959-10
21	33	48.5	461	5	PCT-US93-12464-2
22	33	48.5	554	1	US-08-106-761-2
23	33	48.5	561	2	US-08-532-795-2
24	33	48.5	569	2	US-08-532-795-23
25	33	48.5	569	2	US-08-532-795-29
26	33	48.5	570	2	US-08-532-795-25
27	33	48.5	571	2	US-08-532-795-19

28	33	48.5	571	2	US-08-532-795-21	Sequence 21, Appl
29	33	48.5	574	2	US-08-532-795-27	Sequence 27, Appl
30	33	48.5	614	1	US-08-262-338A-4	Sequence 4, Appl1
31	33	48.5	614	1	US-08-460-114A-4	Sequence 4, Appl1
32	33	48.5	907	4	US-08-989-299-12	Sequence 12, Appl
33	33	48.5	1167	1	US-08-485-568A-6	Sequence 6, Appl1
34	33	48.5	1167	2	US-08-590-554A-6	Sequence 6, Appl1
35	33	48.5	1167	2	US-09-184-223-6	Sequence 6, Appl1
36	33	48.5	1168	1	US-08-620-717A-9	Sequence 9, Appl1
37	32	47.1	10	2	US-08-556-597-168	Sequence 168, App
38	32	47.1	27	1	US-08-318-193-39	Sequence 39, Appl
39	32	47.1	40	1	US-08-318-193-6	Sequence 6, Appl1
40	32	47.1	40	6	5200327-17	Patent No. 5200327
41	32	47.1	118	4	US-09-370-253-18	Sequence 18, Appl
42	32	47.1	152	2	US-08-933-750C-33	Sequence 33, Appl
43	32	47.1	152	4	US-09-234-613-33	Sequence 33, Appl
44	32	47.1	184	2	US-08-602-359A-33	Sequence 33, Appl
45	32	47.1	330	4	US-09-188-930-144	Sequence 144, Appl

ALIGNMENTS

RESULT 1
US-08-439-131A-5
Sequence 5, Application US/08439131A
Patent No. 5512472
GENERAL INFORMATION:
APPLICANT: Lai, Margaret H. K.
APPLICANT: Kirsch, Donald R.
TITLE OF INVENTION: A DNA Sequence Encoding Sterol Delta
Patent No. 5512472
TITLE OF INVENTION: Reductase
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,131A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,347
FILING DATE: 16-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 854-012 (32,141)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 amino acids
TYPE: amino acid
SPRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORGANISM: Schizosaccharomyces pombe
PUBLICATION INFORMATION:
AUTHORS: Shimanuki, M.
AUTHORS: Goebli, M.

AUTHORS: Yanagida, M.
AUTHORS: Toda, T.
TITLE: Fission Yeast stsl+ Gene Encodes a Protein
TITLE: Similar to the Chicken Lamin B Receptor
JOURNAL: Molecular Biology of the Cell
VOLUME: 3
PAGES: 263-273
DATE: 1992
US-08-439-131A-5

Query Match 52.2%; Score 35.5; DB 1; Length 453;
Best Local Similarity 54.5%; Pred. No. 77;
Matches 6; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

OY 2 WPHVLEYLAVG 12
DB 217 WF-ILFFISVG 226

RESULT 2
US-08-440-674-4
Sequence 4, Application US/08440674
Patent No. 5525496
GENERAL INFORMATION:
APPLICANT: Margaret H. Lai
TITLE OF INVENTION: A DNA Sequence Encoding Sterol
Patent No. 5525496
TITLE OF INVENTION: 14
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 Mb diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,674
FILING DATE: May 15, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,347
FILING DATE: August 16, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Alan M. Gordon
REGISTRATION NUMBER: 30637
REFERENCE/DOCKET NUMBER: 854-P0012 Div (32,141)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
FEATURE:
NAME/KEY: Schizosaccharomyces pombe sts gene
PUBLICATION INFORMATION:
AUTHORS: M. Shimanuki, M. Geob, M. Yanagida,
AUTHORS: and T.
AUTHORS: Toda
TITLE: Fission Yeast stsl+ Gene
TITLE: Encodes a

TITLE: Protein Similar to the Chicken Lamin B Receptor
JOURNAL: Molecu-
JOURNAL: lar Biology of the Cell
VOLUME: 3
PAGES: 263-273
PAGES: Sequence set out in Figure 1, page 264
Patent No. 5525496
DATE: 1992
US-08-440-674-4

Query Match 52.2%; Score 35.5; DB 1; Length 453;
Best Local Similarity 54.5%; Pred. No. 77;
Matches 6; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

OY 2 WPHVLEYLAVG 12
DB 217 WF-ILFFISVG 226

RESULT 3
US-08-415-751-36
Sequence 36, Application US/08415751
Patent No. 5643772
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEBCH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUT, JIRI
TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES; DNA
TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
TITLE OF INVENTION: METHODS FOR IMMOTHERAPY AND
TITLE OF INVENTION: DIAGNOSIS AND KIT
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PHILLIPS, MOORE, LEMPIO & FINLEY
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: California
COUNTRY: United States of America
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.5 inch, 1.44 KB storage
COMPUTER: PC
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,751
FILING DATE: 03-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/071,880
FILING DATE: June 1, 1993
APPLICATION NUMBER: 07/891,301
FILING DATE: May 29, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Hana Dolezalova
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.19-2 (HMD)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-1677
TELEFAX: (415) 324-1678
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:

ORGANISM: Cryptosporidium parvum
FEATURE:
NAME/KEY: Positions coded by nonsense codons are
NAME/KEY: Identified as Xaa.
US-08-415-751-36

Query Match 51.5%; Score 35; DB 1; Length 361;
Best Local Similarity 33.3%; Pred. No. 74;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 WPHVLFYLAAG 12
: 11 : 1 : 1 :
DB 267 IMFXTIXYMWIG 278

RESULT 4
US-08-624-125-5
Sequence 5, Application US/08624125
Patent No. 5744341
GENERAL INFORMATION:
APPLICANT: CUNNINGHAM JR., FRANCIS X.
APPLICANT: SUN, ZAIREN
TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,125
FILING DATE: 29-MAR-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: KELLER, STEVEN B.
REGISTRATION NUMBER: 30,073
REFERENCE/DOCKET NUMBER: 2747-063-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-624-125-5

Query Match 50.0%; Score 34; DB 1; Length 162;
Best Local Similarity 45.5%; Pred. No. 47;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 WPHVLFYLAAG 12
: 11 : 1 : 1 :
DB 70 WPHVLFYLAAG 80

RESULT 5
US-08-624-125-8
Sequence 8, Application US/08624125
Patent No. 5744341

GENERAL INFORMATION:
APPLICANT: CUNNINGHAM JR., FRANCIS X.
APPLICANT: SUN, ZAIREN
TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,125
FILING DATE: 29-MAR-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: KELLER, STEVEN B.
REGISTRATION NUMBER: 30,073
REFERENCE/DOCKET NUMBER: 2747-063-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-624-125-8

Query Match 50.0%; Score 34; DB 1; Length 162;
Best Local Similarity 41.7%; Pred. No. 47;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 WPHVLFYLAAG 12
: 11 : 1 : 1 :
DB 69 WPHVLFYLAAG 80

RESULT 6
US-08-663-310-4
Sequence 4, Application US/08663310
Patent No. 5811273
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5811273hiko
APPLICANT: KONO, Keiji
APPLICANT: KAWABARA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1-0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663/310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-663-310-4

Query Match 50.0%; Score 34; DB 2; Length 162;
Best Local Similarity 41.7%; Pred. No. 47;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPHVLYAVG-12
DB 69 IWPVLMWIALG 80

RESULT 7
US-08-663-310-11
Sequence 11, Application US/08663310
Patent No. 5811273
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5811273hiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIMAWA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-663-310-11

Query Match 50.0%; Score 34; DB 2; Length 162;
Best Local Similarity 45.5%; Pred. No. 47;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 WPHVLYAVG 12
DB 70 WWPVLMWIALG 80

RESULT 8
US-09-006-491-4
Sequence 4, Application US/09006491
Patent No. 5972690
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5972690hiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIMAWA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-006-491-4

Query Match 50.0%; Score 34; DB 2; Length 162;
Best Local Similarity 41.7%; Pred. No. 47;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 WPFVLEFYLAG 12
DB 69 IMAVPLMWITLG 80

RESULT 9
US-09-006-491-11
Sequence 11, Application US/09006491
Patent No. 5972690
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5972690hiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIWARA, Susumu
APPLICANT: YOKOYAMA, Akihiro
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF INVENTION: XANTHOPHYLLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-006-491-11

Query Match 50.0%; Score 34; DB 2; Length 162;
Best Local Similarity 45.5%; Pred. No. 47;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 WPFVLEFYLAG 12
DB 70 IMAVPLMWITLG 80

RESULT 10
US-09-335-919-4
Sequence 4, Application US/09335919
Patent No. 6150130
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 6150130hiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIWARA, Susumu
APPLICANT: YOKOYAMA, Akihiro
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF INVENTION: XANTHOPHYLLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,919
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-335-919-4

ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-49

Query Match 50.0%; Score 34; DB 1; Length 340;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWFHVLFF 8
11:111
Db 61 WYVGLFF 68

RESULT 14
PCT-US93-08528-49
Sequence 49, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033

REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-49

Query Match 50.0%; Score 34; DB 5; Length 340;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWFHVLFF 8
11:111
Db 61 WYVGLFF 68

RESULT 15
US-07-937-609-20
Sequence 20, Application US/07937609
Patent No. 5319073
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor

Tue Jun 5 07:09:50 2001

us-09-284-107-29.rai

Page 8

US-07-937-609-20

Query Match

50.0%; Score 34; DB 1; Length 372;

Best Local Similarity 62.5%; Pred. NO. 1.2e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VWFHLEFY-8

Db 97 VWFHLEFY 104

Search completed: June 4, 2001, 12:14:33
job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:15:46 ; Search time 65.22 Seconds
(without alignments)
12.645 Million cell updates/sec

Title: US-09-284-107-29

Perfect score: 68

Sequence: 1 VWFHVLFLAVG 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR_67:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	280	2	155577
2	68	100.0	344	2	A41357
3	68	100.0	374	1	A39878
4	64	94.1	336	2	I48471
5	64	94.1	404	2	A46480
6	43.5	64.0	394	2	D65167
7	41	60.3	153	2	C72420
8	41	60.3	340	2	T20102
9	40	60.3	1852	2	JC5546
10	41	58.8	88	2	T17560
11	40	58.8	448	2	A69747
12	39	57.4	421	2	SA0819
13	38	55.9	84	2	F69202
14	38	55.9	266	2	F69974
15	38	55.9	311	2	T11017
16	38	55.9	314	2	A71157
17	38	55.9	343	2	T06057
18	38	55.9	354	2	S66923
19	38	55.9	357	2	T07418
20	38	55.9	610	2	S19461
21	38	55.9	1869	2	A59290
22	37	54.4	194	2	T11071
23	37	54.4	210	1	S74656
24	37	54.4	267	2	I72882
25	37	54.4	267	2	I56110
26	37	54.4	267	2	A35902
27	37	54.4	267	2	C70871
28	37	54.4	375	2	A81227
29	37	54.4	375	2	F81999

ALIGNMENTS

30	37	54.4	405	2	E81946	hypothetical prote
31	37	54.4	407	2	H81160	hypothetical prote
32	37	54.4	473	2	S51256	probable membrane
33	37	54.4	621	2	T48646	phytoene dehydrog
34	37	54.4	776	2	I48317	cyclin F - mouse
35	36	52.9	82	2	T51909	related to cytochr
36	36	52.9	123	2	S63055	probable membrane
37	36	52.9	242	2	E81315	CDDdiacylglycerol-
38	36	52.9	252	2	G75485	hypothetical prote
39	36	52.9	254	2	T20882	hypothetical prote
40	36	52.9	366	1	XNBYUG	UDPglucose--hexose
41	36	52.9	368	2	I52863	myosin heavy chain
42	36	52.9	370	1	XNVKUD	UDPglucose--hexose
43	36	52.9	483	2	T16443	hypothetical prote
44	36	52.9	515	2	H84013	hypothetical prote
45	36	52.9	555	2	H83043	hypothetical prote

RESULT 1

155577
Fc gamma (Igc) receptor I-B splice form 1 precursor - human
N:Alternate names: CD64
M:Contains: Fc-gamma (Igc) receptor I-B splice form 2
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: 155577; I70303
R:Portes, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
J Clin Invest. 90, 2102-2109, 1992
A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
A:Reference number: 155577; MUID:93053454
A:Accession: 155577
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-280 <RES>
A:Cross-references: GB:L03419; NID:q182460; PIDN:AAA35825.1; PID:q292023
A:Note: splice form B1
A:Accession: I70303
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-10,103-153, 'A',155-280 <RES>
A:Cross-references: GB:L03420; NID:q182461; PIDN:AAA35826.1; PID:q292024
A:Experimental source: mononuclear cells
A:Note: splice form B2
C:Comment: This receptor does not bind monomeric Igc with high affinity.
C:Genetics:
A:Gene: GDB:FCGR1B; CD64
A:Cross-references: GDB:135923; OMIM:601502
A:Map position: 1p12-1p12
A:Superfamily: Fc gamma receptor I; immunoglobulin homology
C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembran
F.117-170/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 68; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWFHVLFLAVG 12
DB 195 VWFHVLFLAVG 206

RESULT 2
A1357
Fc gamma (Igc) receptor I (high affinity) form b - human
N:Alternate names: CD64
C:Species: Homo sapiens (man)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jan-2000
C:Accession: A1357; S03019
R:Allen, J.M.; Seed, B.

Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: A41357; MUID:89100284
 A:Accession: A41357
 A:Molecule type: mRNA
 A:Residues: 1-344 <AL1>
 A:Cross-references: GB:X14355; GB:M21090; NID:g31333; PIDN:CAA32536.1; PID:g31334
 R:Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (Fcr)
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-344 <ALL2>
 A:Cross-references: EMBL:X14355; NID:g31333; PIDN:CAA32536.1; PID:g31334
 A>Note: the authors translated the codon ACT for residue 25 as Ala
 C:Superfamily: Fc gamma receptor I; Immunoglobulin homology
 C:Keywords: alternative splicing; immunoglobulin receptor; transmembrane protein
 F:117-170/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 68; DB 2; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.00019;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHLYFYLAAG 12
 |||||
 DB (289 VWFHLYFYLAAG 300

RESULT 3
 A39878
 Fc gamma (IgG) receptor I-A (high affinity) precursor - human
 N:Alternate names: CD64
 C:Species: Homo sapiens (man)
 C>Date: 30-Dec-1991 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
 C:Accession: A39878; I70304; B41357; S03018; I57525
 R:van de Winkel, J.G.J.; Ernst, L.K.; Anderson, C.L.; Chiu, I.M.
 J. Biol. Chem. 266, 13449-13455, 1991
 A:Title: Gene organization of the human high affinity receptor for IgG, Fc gamma (CD64)
 A:Reference number: A39878; MUID:91302383
 A:Accession: A39878
 A:Molecule type: DNA
 A:Residues: 1-374 <VAN>
 A:Cross-references: GB:M63830; GB:M63835; NID:g180277; PIDN:AAA35678.1; PID:g180279
 R:Porces, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, R.P.
 J. Clin. Invest. 90, 2102-2109, 1992
 A:Title: Novel Fc gamma receptor I family gene products in human mononuclear cells.
 A:Reference number: I55577; MUID:93055454
 A:Accession: I70304
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-374 <RES>
 A:Cross-references: GB:I03418; NID:g184840; PIDN:AAA36049.1; PID:g292169
 R:Allen, J.M.; Seed, B.
 Science 243, 378-381, 1989
 A:Title: Isolation and expression of functional high-affinity Fc receptor complementary
 A:Reference number: A41357; MUID:89100284
 A:Accession: B41357
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'W', '339-374' <AL1>
 A:Cross-references: GB:X14356; GB:M21091; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R:Allen, J.M.; Seed, B.
 Nucleic Acids Res. 16, 11824, 1988
 A:Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (Fcr)
 A:Reference number: S03018; MUID:89098339
 A:Accession: S03018
 A:Molecule type: mRNA
 A:Residues: 1-24, 'S', '26-337', 'W', '339-374' <AL2>
 A:Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332
 R:Peres, C.; Wietzeblin, J.; Benesch, P.D.
 Mol. Cell. Biol. 13, 2182-2192, 1993
 A:Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma int

15M.
 A:Reference number: I57525; MUID:93204964
 A:Accession: I57525
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8 <RE2>
 A:Cross-references: GB:S57204; NID:g298692; PIDN:AAD13887.1; PID:g4261587
 C:Genetics: GDB:FCGR1A; CD64
 A:Cross-references: GDB:135911; OMIM:146760
 A:Map position: 1q21-1q21
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembran
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-292/Domain: extracellular #status predicted <EXT>
 F:117-170/Domain: immunoglobulin homology <IMM2>
 F:293-313/Domain: transmembrane #status predicted <TM2>
 F:59,78,152,159,163,195,240/Binding site: carbohydrate (asn) (covalent) #status predi

Query Match 100.0%; Score 68; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.0002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHLYFYLAAG 12
 |||||
 DB 289 VWFHLYFYLAAG 300

RESULT 4
 I48471
 Fc gamma (IgG) receptor high affinity - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
 C:Accession: I48471
 R:Pirns, J.B.; Todd, J.A.; Rodrigues, N.R.; Ghosh, S.; Hogarth, P.M.; Wicker, L.S.; G
 Science 260, 695-698, 1993
 A:Title: Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for
 A:Reference number: I48471; MUID:93242399
 A:Accession: I48471
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-336 <RES>
 A:Cross-references: EMBL:X70980; NID:g311748; PIDN:CAA50311.1; PID:g311749
 C:Superfamily: Fc gamma receptor I; immunoglobulin homology
 C:Keywords: immunoglobulin receptor
 F:128-180/Domain: immunoglobulin homology <IMM>

Query Match 94.1%; Score 64; DB 2; Length 336;
 Best Local Similarity 83.3%; Pred. No. 0.0009;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHLYFYLAAG 12
 |||||
 DB 303 VWFHLYFYLAAG 314

RESULT 5
 A46480
 Fc gamma (IgG) receptor high affinity - mouse
 N:Alternate names: high affinity IgG receptor
 C:Species: Mus musculus (house mouse)
 C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
 C:Accession: A46480; A43511
 R:Osman, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.
 J. Immunol. 148, 1570-1575, 1992
 A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma R1 a
 A:Reference number: A46480; MUID:92166399
 A:Accession: A46480
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-404 <OSM>

A:Note: sequence extracted from NCBI backbone (NCBIN:85205, NCBIN:85208, NCBIN:85212, NC
 B:sears, D.W.; Osman, N.; Tate, B.; McKenzie, I.F.C.; Hogarth, P.M.
 J. Immunol. 144, 371-378, 1990
 A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for IgG
 A:Reference number: A43511; MUID:90111035
 A:Accession: A43511
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-404 <SER>
 A:Cross-references: GB:M31314; NID:g200752; PIDN:AAA40056.1; PID:g200753
 C:Superfamily: Fc gamma receptor II; immunoglobulin homology
 C:Keywords: immunoglobulin receptor; transmembrane protein
 F:127-179/Domain: immunoglobulin homology <IMM>

Query Match 94.1%; Score 64; DB 2; Length 404;
 Best Local Similarity 83.3%; Pred. No. 0.0011;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHLYFLAVG 12
 DB 298 VWFHLYFLVSG 309

RESULT 6
 D65167
 Probable membrane protein yick - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 17-Mar-2000
 C:Accession: D65167
 R:Baltner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
 A.: Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: D65167
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-394 <BLAT>
 A:Cross-references: GB:AE000443; GB:U00096; NID:g2367255; PIDN:AAC76682.1; PID:g1790091;
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: yick
 C:Superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein MTH78
 C:Keywords: transmembrane protein
 F:16-32/Domain: transmembrane #status predicted <TM1>
 F:50-66/Domain: transmembrane #status predicted <TM2>
 F:82-98/Domain: transmembrane #status predicted <TM3>
 F:105-121/Domain: transmembrane #status predicted <TM4>
 F:116-192/Domain: transmembrane #status predicted <TM5>
 F:258-274/Domain: transmembrane #status predicted <TM6>
 F:284-300/Domain: transmembrane #status predicted <TM7>
 F:314-330/Domain: transmembrane #status predicted <TM8>
 F:375-391/Domain: transmembrane #status predicted <TM9>

Query Match 64.0%; Score 43.5; DB 2; Length 394;
 Best Local Similarity 53.8%; Pred. No. 3.8;
 Matches 7; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

OY 1 VW-FHLYFLAVG 12
 DB 367 IWSYHALFWLAIG 379

RESULT 7
 C72420
 Hypothetical protein TM0089 - Thermotoga maritima (strain MSB)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: C72420
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316
 A:Accession: C72420
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-153 <ARN>
 A:Cross-references: GB:AE001695; GB:AE000512; NID:g4980569; PIDN:AAD35183.1; PID:g498
 A:Experimental source: strain MSB
 C:Genetics:
 A:Gene: TM0089
 C:Superfamily: Thermotoga maritima hypothetical protein TM0089

Query Match 60.3%; Score 41; DB 2; Length 153;
 Best Local Similarity 66.7%; Pred. No. 4.1;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 VWFHLYFLAVG 12
 DB 15 VWSFVLYFLVAG 26

RESULT 8
 T20102
 Hypothetical protein C50C10.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T20102
 R:McMurray, A.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19224
 A:Accession: T20102
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-340 <MIT>
 A:Cross-references: EMBL:Z72505; PIDN:CAA96610.1; GSPDB:GN00023; CESP:C50C10.4
 A:Experimental source: clone C50C10
 C:Genetics:
 A:Gene: CESP:C50C10.4
 A:Map position: 5
 A:Introns: 79/3; 149/3; 272/3

Query Match 60.3%; Score 41; DB 2; Length 340;
 Best Local Similarity 54.5%; Pred. No. 8.9;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 WPHVLYFLAVG 12
 DB 282 WYHLYFLRPIG 292

RESULT 9
 JC5546
 Chitin synthase (EC 2.4.1.16) asma - Emericella nidulans
 N:Alternate names: chitin-UDP N-acetylglucosaminyltransferase
 C:Species: Emericella nidulans, Aspergillus nidulans
 C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 20-Aug-1999
 C:Accession: JC5546
 R:Fujiwara, M.; Horinuchi, H.; Ohta, A.; Takagi, M.
 Biochem. Biophys. Res. Commun. 236, 75-78, 1997
 A:Title: A novel fungal gene encoding chitin synthase with a myosin motor-like domain
 A:Reference number: JC5546; MUID:97366599
 A:Accession: JC5546
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1852 <EU>
 A:Cross-references: DDBJ:AB000125; NID:g2308976; PID:d1022569; PID:g2308977
 C:Comment: This enzyme is a membrane-bound protein. It is involved in the catalytic p
 C:Genetics:
 A:Gene: asma

C:Keywords: ATP; glycosyltransferase; hexosyltransferase
F:102-110/domain: ATP-binding #status predicted <ATP>

Query Match 60.3%; Score 41; DB 2; Length 1852;
Best Local Similarity 70.0%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 FHVLEFLAVG 12
||||:|:|
Db 198 FHVLYLAVG 207

RESULT 10

T17560
hypothetical protein a70L - Chlorella virus PBCY-1
C:Species: Chlorella virus PBCY-1
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Accession: T17560
A:Reference number: Z18806
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-88 <GRA>
A:Cross-references: EMBL:U02580; NID:G4028896; PIDN:AAC96438.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Gene: a70L

Query Match 58.8%; Score 40; DB 2; Length 88;
Best Local Similarity 66.7%; Pred. No. 3.5;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VWHVLEFL 9
:|:|:|:|
Db 21 IMFHLEFL 29

RESULT 11

ABC transporter (permease) homolog ybdb - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: A69747
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehlich, S.D.; Emmertson, P.T.; Ehtian, K.D.; Erington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akuch, M.; Tanakoshi, A.; Tanaka, T.; Tepstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A.; Authors: Yoshikawa, H.F.; Zumbach, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033
A:Accession: A69747
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-448 <KUN>
A:Cross-references: GB:Z99105; GB:AL009126; NID:G2632457; PIDN:CAB11990.1; PID:el182148;
A:Experimental source: strain 168
C:Genetics:
A:Gene: ybdb

Query Match 58.8%; Score 40; DB 2; Length 448;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 WFWVLEFLAV 11
||:|:|:|
Db 296 WFWLLEFLIV 305

RESULT 12

S40819
hypothetical 46.3k protein (glnA-fdhe intergenic region) - Escherichia coli
N:Alternate names: hypothetical protein o421
C:Species: Escherichia coli
C:Date: 19-May-1994 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999
R:Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 21, 3391-3398, 1993
A:Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region fro
A:Reference number: S40802; MUID:93347969
A:Accession: S40819
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-421 <PLU>
A:Cross-references: EMBL:L19201; NID:G304961; PIDN:AAB03008.1; PID:G304979
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: E65192
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-421 <BLA>
A:Cross-references: GB:AE000463; GB:U00096; NID:G2367320; PIDN:AAC76871.1; PID:G17903
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yjhn
C:Superfamily: Escherichia coli hypothetical protein (glnA-fdhe intergenic region)

Query Match 57.4%; Score 39; DB 2; Length 421;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 WFWVLEFLAVG 12
||:|:|:|
Db 171 WFWLLEFLIVAVG 182

RESULT 13

F69202
hypothetical protein MTH768 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: F69202
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
Qiu, D.; Spadofora, R.; Vitale, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
K.I.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514
A:Accession: F69202
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-84 <MTH>
A:Cross-references: GB:AE000855; GB:AE000666; NID:G2621852; PIDN:AAB85271.1; PID:G262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH768

Query Match 55.9%; Score 38; DB 2; Length 84;
Best Local Similarity 36.4%; Pred. No. 7.5;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 VWFHVLFLYAV 11
Db 4 IMFYIALFLAI 14

RESULT 14

F69974

formate dehydrogenase homolog yrhg - Bacillus subtilis
C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: F69974

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gallier
tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hult, M.F.
Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelid
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scofione, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:80404033

A:Accession: F69974

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-266 <RUN>

A:Cross-references: GB:299117; GB:AL009126; NID:g2634966; PIDN:CAB14662.1; PID:g2635166

A:Experimental source: strain 168

C:Genetics:

A:Gene: yrhg

C:Superfamily: formate dehydrogenase foch

Query Match 55.9%; Score 38; DB 2; Length 266;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 VWFHVLFLYAV 12
Db 189 IMFYIALFLAI 200

RESULT 15

T11017

MLC1536.34 protein - Mycobacterium leprae
C:Species: Mycobacterium leprae

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Nov-2000
C:Accession: T11017; S72779

R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997

A:Reference number: Z16918

A:Accession: T11017

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-311 <PAR>

A:Cross-references: EMBL:Z29125; NID:g2398683; PIDN:CAB16178.1; PID:g2398713

R:Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid B1496.

A:Reference number: S72693

A:Accession: S72779

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 'M', 90-311 <SMT>

A:Cross-references: EMBL:U00013; NID:g466868; PIDN:AAA17144.1; PID:g466899
C:Genetics:
A:Gene: MLC1536.34
A:Start codon: GTG

Query Match 55.9%; Score 38; DB 2; Length 311;
Best Local Similarity 58.3%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 VWFHVLFLYAV 12
Db 151 VMLAVLLYKVG 162

Search completed: June 4, 2001, 12:15:47
Job time: 280 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:24:09 ; Search time 37.56 Seconds

(Without alignments)
10,944 Million cell updates/sec

Title: US-09-284-107-29

Perfect score: 68
Sequence: 1 VWFHVLFLYAVG 12

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	374	1	FCG1_HUMAN
2	64	94.1	404	1	FCG1_MOUSE
3	43.5	64.0	394	1	SEPC_ECOLI
4	39	57.4	421	1	Y1HN_ECOLI
5	38	55.9	172	1	AD04_HUMAN
6	38	55.9	266	1	YRNG_BACSU
7	38	55.9	610	1	ARE1_YEAST
8	37	54.4	210	1	YH90_SYNY3
9	37	54.4	267	1	FCG3_RAT
10	37	54.4	398	1	CTM4_MOUSE
11	37	54.4	621	1	CRT1_CERNC
12	37	54.4	776	1	CG2F_MOUSE
13	36	52.9	123	1	YNLA_YEAST
14	36	52.9	274	1	FDHC_METTE
15	36	52.9	365	1	GAL7_YEAST
16	36	52.9	370	1	GAL7_KLULA
17	36	52.9	387	1	U133_RCMVA
18	36	52.9	873	1	COX1_ACACA
19	36	52.9	919	1	SYL_THEMA
20	36	52.9	1972	1	MYST_RABIT
21	36	52.9	1978	1	MYST_CHICK
22	35	52.2	453	1	STSL_SCHPO
23	35	52.2	629	1	Y252_RICPR
24	35	51.5	185	1	VNS2_MOUSE
25	35	51.5	287	1	MAUN_METEX
26	35	51.5	349	1	ALFD_PEA
27	35	51.5	356	1	ALFC_PEA
28	35	51.5	429	1	SECY_AQUAE
29	35	51.5	439	1	SECY_SYNP7
30	35	51.5	486	1	HHIR_RAT
31	35	51.5	499	1	C1W5_HUMAN
32	35	51.5	547	1	STW_BUCAT
33	35	51.5	547	1	STW_BUCAT

ALIGNMENTS

RESULT	1	FCG1_HUMAN	STANDARD	PRT	374 AA
AC	P12314; P12315;				
DT	01-OCT-1989 (Rel. 12, Created)				
DT	01-OCT-1989 (Rel. 12, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I) (CD64).				
GN	FCGRIA OR FCGRI OR FCGI OR IGFRI.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89098339; PubMed=2974947;				
RA	Allen J.M., Seed B.;				
RT	*Nucleotide sequence of three cDNAs for the human high affinity Fc				
RL	receptor (FcRI)."				
RN	Nucleic Acids Res. 16:11824-11824(1988).				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89100284; PubMed=2911749;				
RA	Allen J.M., Seed B.;				
RT	*Isolation and expression of functional high-affinity Fc receptor				
RL	complementary DNAs."				
RT	Science 243:378-381(1989).				
CC	- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH				
CC	AFFINITY RECEPTOR.				
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B: ARE				
CC	PRODUCED BY ALTERNATIVE SPLICING.				
CC	- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.				
CC	- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.				
CC	- DATABASE: NAME=PROW; NOTE=CD guide C64 entry;				
CC	WWW="http://www.ncbi.nlm.nih.gov/prov/cd/c64.htm"				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/				
CC	or send an email to license@isb-sdb.ch).				
DR	EMBL; X14356; CAA32537.1; -				
DR	EMBL; X14355; CAA32536.1; -				
DR	PIR; S03018; S03018.				
DR	PIR; S03019; S03019.				
DR	PIR; A41357; A41357.				
DR	PIR; B41357; B41357.				
DR	HSSP; P12319; 1ALT.				
DR	MIM; 146760; -				
DR	InterPro: IPR003006; -				
DR	Pfam: PF00047; 19; 3.				

34	35	51.5	555	1	NUSM_CANPA
35	580	1	P66_MYCHR		P48919 candida par
36	35	51.5	940	1	SYL_BUCAT
37	34.5	50.7	36	1	Y16L_BPT4
38	34.5	50.7	94	1	TRA0_ECOLI
39	34.5	50.7	374	1	E13B_HEVBR
40	34	50.0	110	1	HIS2_AZOCB
41	34	50.0	162	1	CRT2_AZGRU
42	34	50.0	162	1	CRT2_AICSP
43	34	50.0	219	1	YN82_YEAST
44	34	50.0	261	1	FCG3_MOUSE
45	34	50.0	280	1	FDHC_METFO
					P48919 candida par
					P15362 mycoplasma
					P57249 buchnera ap
					P39244 bacterioph
					P18033 escherichia
					P52407 hevea brasl
					Q43926 azotobacter
					P54973 agrobacteri
					Q44262 alcaligenes
					P53747 saccharomyc
					P08508 mus musculu
					P35839 methanobact

KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain; Alternative splicing; Polymorphism;
 FT SIGNAL 21
 FT CHAIN 1
 FT 22 374
 FT DOMAIN 22 292
 FT TRANSMEM 293 313
 FT DOMAIN 314 374
 FT DOMAIN 36 92
 FT DOMAIN 117 175
 FT DOMAIN 205 267
 FT DISULFID 43 85
 FT DISULFID 124 168
 FT DISULFID 212 260
 FT CARBOHYD 59 59
 FT CARBOHYD 78 78
 FT CARBOHYD 152 152
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 195 195
 FT CARBOHYD 240 240
 FT VARSPLIC 333 374
 FT VARIANT 26 26
 FT VARIANT 59 59
 FT VARIANT 59 59
 FT CONFLICT 25 25
 FT SEQUENCE 374 AA; 42605 MW; 2C2AA8103ECF16E6 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 68; DB 1; Length 374;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWFHLYFLAVG 12
 DB 289 VWFHLYFLAVG 300

RESULT 2
 ID FCGL_MOUSE STANDARD; PRT; 404 AA.
 AC P26151;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR I PRECURSOR (FC-GAMMA RI) (FCRI) (IGG FC RECEPTOR I).
 DE FCGR1 OR FCGL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90111035; PubMed=2136886;
 RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
 RT "Molecular cloning and expression of the mouse high affinity Fc receptor for IgG."
 RL J. Immunol. 144:371-378 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92166399; PubMed=1531670;
 RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
 RT "Structure and mapping of the gene encoding mouse high affinity Fc gamma RI and chromosomal location of the human Fc gamma RI gene."
 RL J. Immunol. 148:1570-1575 (1992).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN GAMMA. HIGH AFFINITY RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: M31314; AAA40056.1;
 CC PIR: A43511; A43511.
 CC PIR: A46480; A46480.
 CC HSSP: P12319; 1ALT.
 CC MGD: MGI:95498; Fcgr1.
 CC InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 3.
 DR
 KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1
 FT CHAIN 25 404
 FT DOMAIN 25 297
 FT TRANSMEM 298 320
 FT DOMAIN 321 404
 FT DOMAIN 46 66
 FT DOMAIN 67 124
 FT DOMAIN 154 216
 FT CARBOHYD 28 28
 FT CARBOHYD 48 48
 FT CARBOHYD 69 69
 FT CARBOHYD 168 168
 FT CARBOHYD 249 249
 FT SEQUENCE 404 AA; 44887 MW; 1CAF0033842767E7 CRC64;

Query Match
 Best Local Similarity 94.1%; Score 64; DB 1; Length 404;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWFHLYFLAVG 12
 DB 298 VWFHLYFLAVG 309

RESULT 3
 ID SETC_ECOLI STANDARD; PRT; 394 AA.
 AC P31436;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SUGAR EFFLUX TRANSPORTER C.
 DE SETC.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=K12 / MG1655;
 RX MEDLINE=93315143; PubMed=7666882;
 RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
 RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli genome: organizational symmetry around the origin of replication."
 RL Genomics 16:551-561 (1993).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=99226230; PubMed=10209755;
 RA Liu J.Y., Miller P.F., Gosink M., Olson E.R.;
 RT "The identification of a new family of sugar efflux pumps in Escherichia coli."
 RL Mol. Microbiol. 31:1845-1851 (1999).

RN [3]
 RP CHARACTERIZATION. PubMed-10438463;
 RX MEDLINE-99367417; Willard J., Olson E.R.;
 RA Liu J.Y., Miller P.F., Willard J., Olson E.R.;
 RT "Functional and biochemical characterization of Escherichia coli sugar
 efflux transporters.";
 RL J. Biol. Chem. 274:22977-22984(1999).
 CC -1- FUNCTION: INVOLVED IN THE EFFLUX OF SUGARS. THE PHYSIOLOGICAL ROLE
 CC MAY BE THE DETOXIFICATION OF NON-METABOLIZABLE SUGAR ANALOGS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE SET FAMILY OF TRANSPORTERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L10328; AAA62011.1;
 DR EMBL: AE000443; AAC76682.1;
 DR Ecocore: EG11687; setc.
 KW Transport: Sugar transport; Transmembrane; Inner membrane.
 FT TRANSMEM 11 31 POTENTIAL.
 FT TRANSMEM 50 70 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 151 171 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 225 245 POTENTIAL.
 FT TRANSMEM 254 274 POTENTIAL.
 FT TRANSMEM 284 304 POTENTIAL.
 FT TRANSMEM 311 331 POTENTIAL.
 FT TRANSMEM 371 391 POTENTIAL.
 SQ SEQUENCE 394 AA; 43493 MW; 310838BDD04F92E CRC64;

 Query Match 64.0%; Score 43.5; DB 1; Length 394;
 Best Local Similarity 53.8%; Pred. No. 4.2;
 Matches 7; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

 OY 1 VW-FHYVLYAVG 12
 Db 367 IMVSYALFVLAIG 379

 RESULT 4
 YIHN_ECOLI STANDARD; PRT; 421 AA.
 ID YIHN_ECOLI
 AC P32135;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYDROTHERMAL 46.3 KDA PROTEIN IN GLNA-RBN INTERGENIC REGION (0421).
 GN YIHN.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBL_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-93347969; PubMed-8346018;
 RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
 RT region from 87.2 to 89.2 minutes.";
 RL Nucleic Acids Res. 21:3391-3398(1993).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: TO E.COLI VOCE.
 CC -1- SIMILARITY: SOME, TO M.GENITALIUM MG294 AND TO M.JANNASCHII
 CC M03342.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L19201; AAB03008.1;
 DR EMBL: AE000463; AAC76871.1;
 DR PIR: S40819; S40819.
 DR Ecocore: EG11840; yihN.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 45 65 POTENTIAL.
 FT TRANSMEM 88 108 POTENTIAL.
 FT TRANSMEM 147 167 POTENTIAL.
 FT TRANSMEM 169 189 POTENTIAL.
 FT TRANSMEM 221 241 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 289 309 POTENTIAL.
 FT TRANSMEM 310 330 POTENTIAL.
 FT TRANSMEM 355 375 POTENTIAL.
 FT TRANSMEM 386 406 POTENTIAL.
 SQ SEQUENCE 421 AA; 46335 MW; 4FD9FAA33ABEB46 CRC64;

 Query Match 57.4%; Score 39; DB 1; Length 421;
 Best Local Similarity 50.0%; Pred. No. 23;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

 OY 1 VWFHYVLYAVG 12
 Db 171 IMVSYALFVLAIG 182

 RESULT 5
 AD04_HUMAN STANDARD; PRT; 172 AA.
 ID AD04_HUMAN
 AC O9Y3D8;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROTEIN AD-004 (PROTEIN CGI-137).
 GN AD004.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBL_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Adrenal gland;
 RX MEDLINE-20402571; PubMed-10931946;
 RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
 RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Kong R., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 RT axis and full-length cDNA cloning.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-20272150; PubMed-10810093;
 RX Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;
 RT "Identification of novel human genes evolutionarily conserved in
 RT Caenorhabditis elegans by comparative proteomics.";
 RL Genome Res. 10:703-713(2000).
 CC -1- SIMILARITY: BELONGS TO THE UPF0101 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF110777; AAF14860.1; -
DR EMBL: AF151895; AAD34132.1; -
DR InterPro: IPR000623; -
DR PRINTS: PRO1100; SHIKIMTKINASE.
KW ATP-binding.
FT NP BIND 10 17 ATP (POTENTIAL)
SQ SEQUENCE 172 AA: 20061 MW: 335623185B434301 CRC64:

Query Match 55.9%; Score 38; DB 1; Length 172;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 VWHVLFYL 9
DB 88 VWHVLFVL 95

RESULT 6

YRHG_BACSU STANDARD; PRT; 266 AA.

ID YRHG_BACSU STANDARD; PRT; 266 AA.

AC 005399;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE HYPOTHETICAL 28.5 KDA PROTEIN IN SIGV-GREA INTERGENIC REGION.

GN YRHG.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA MEDLINE=97455479; PubMed=9308178;

RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,

RA Duesterhoeft A., Emlich S.D.;

RT "Sequence of the Bacillus subtilis genome region in the vicinity of

RT the lev operon reveals two new extracytoplasmic function RNA

RT polymerase sigma factors Sigv and Sigz."

RL Microbiology 143:2939-2943(1997).

CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC -1 SIMILARITY: BELONGS TO THE FMT FAMILY OF TRANSPORTERS (TC 2.44).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U93874; AAB0864.1; -

DR EMBL: Z99117; CAB14662.1; -

DR Subtilist; BG12296; yrhg.

DR InterPro: IPR000292; -

DR Pfam: PF01226; Form_Ntr trans; 1.

DR PROSITE: PS01005; FORMATE_NITRATE_TP_1; 1.

DR PROSITE: PS01006; FORMATE_NITRATE_TP_2; 1.

KW Hypothetical protein; Transmembrane; Transport.

FT TRANSMEM 25 45 POTENTIAL.

FT TRANSMEM 64 84 POTENTIAL.

FT TRANSMEM 111 131 POTENTIAL.

FT TRANSMEM 158 178 POTENTIAL.

FT TRANSMEM 186 206 POTENTIAL.

FT TRANSMEM 209 229 POTENTIAL.

FT TRANSMEM 230 250 POTENTIAL.

FT TRANSMEM 266 AA: 28488 MW: 64537664FC7C7147 CRC64;

SQ SEQUENCE

Query Match 55.9%; Score 38; DB 1; Length 266;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 VWHVLFYLANG 12
DB 189 IWFPIMAFVAIG 200

RESULT 7

AREL_YEAST STANDARD; PRT; 610 AA.

ID AREL_YEAST STANDARD; PRT; 610 AA.

AC P25628;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE STEROL O-ACYLTRANSFERASE 1 (EC 2.3.1.26) (STEROL-ESTER SYNTHASE 1).

GN AREL OR SAT2 OR YCR048W OR YCR48W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RA Grivell L.A., de Haan M., Maat M.J.;

RA Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE OF 1-328 FROM N.A.

RA Bolotin-Fukuhara M., Buhler J.-M., Daignan-Fornier B., Dolra C.,

RA Franchingues-Galliard M.-C.;

RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.

RP CHARACTERIZATION.

RX MEDLINE=96243137; PubMed=8650549;

RA Yang H., Bard M., Bruner D.A., Gleeson A., Deckelbaum R.J.,

RA Aljinovic G., Pohl T.M., Rothstein R., Sturley S.L.;

RT "Sterol esterification in yeast: a two-gene process."

RL Science 272:1353-1356(1996).

CC -1 CATALYTIC ACTIVITY: ACYL-COA + CHOLESTEROL = COA + CHOLESTEROL

CC ESTER.

CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC

CC RETICULUM.

CC -1 SIMILARITY: TO OTHER STEROL O-ACYLTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X59720; CAA42296.1; -

DR PIR: S19461; S19461.

DR SGD: S0000644; AREL.

DR InterPro: IPR002688; -

DR Pfam: PF01800; ACAT; 1.

KW Transferase; Transmembrane; Endoplasmic reticulum; Acyltransferase.

FT TRANSMEM 182 202 POTENTIAL.

FT TRANSMEM 229 249 POTENTIAL.

FT TRANSMEM 264 284 POTENTIAL.

FT TRANSMEM 371 391 POTENTIAL.

FT TRANSMEM 409 429 POTENTIAL.

FT TRANSMEM 535 555 POTENTIAL.

FT TRANSMEM 590 610 POTENTIAL.

FT TRANSMEM 610 AA: 71613 MW: FF72EFB9238B2205 CRC64;

SQ SEQUENCE

Query Match 55.9%; Score 38; DB 1; Length 610;
Best Local Similarity 54.5%; Pred. No. 46;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 WPHLYFLANG 12
 DB 450 WPHLYFLANG 460

RESULT 8
 ID YH90_SYNY3 STANDARD: PRT: 210 AA.
 AC P2793;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HYPOTHEORETICAL 24.1 KDA PROTEIN SLR1790.
 GN SLR1790.
 OS Synchocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 OX NCBI_Taxid=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE UPF0093 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90900; BAA16808.1;
 KM Hypothetical protein; Transmembrane.
 FT TRANSMEM 22 42 POTENTIAL.
 FT TRANSMEM 74 94 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 SQ SEQUENCE 210 AA; 24063 MW; 137CF30151F75E2 CRC64;

Query Match 54.48; Score 37; DB 1; Length 210;
 Best Local Similarity 77.88; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WPHLYFL 9
 DB .33 WPHLYFL 41

RESULT 9
 ID FC3_RAT STANDARD: PRT: 267 AA.
 AC P27645; 063204; 004798;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION RECEPTOR III PRECURSOR
 DE (IGG FC RECEPTOR III) (FC-GAMMA RIIT) (FCRIIT).
 GN FCGR3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]

RP SEQUENCE FROM N.A. (ISOFORM B; C AND D).
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=91250730; PubMed=1710249;
 RA Farber D.L., Sears D.W.;
 RT "Rat CD16 is defined by a family of class III Fc gamma receptors
 RT requiring co-expression of heteroprotein subunits.";
 RL J. Immunol. 146:4352-4361(1991).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A).
 RC TISSUE=Natural killer cells;
 RX MEDLINE=90239026; PubMed=1692135;
 RA Zeger D.L., Hogarth P.M., Sears D.W.;
 RT "Characterization and expression of an Fc gamma receptor cDNA cloned
 RT from rat natural killer cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3425-3429(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM H).
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
 RX MEDLINE=93246550; PubMed=8482840;
 RA Farber D.L., Giorda R., Nettleton M.Y., Trucco M., Kochan J.P.,
 RA Sears D.W.;
 RT "Rat class III Fc gamma receptor isoforms differ in IgG subclass-
 RT binding specificity and fail to associate productively with rat CD3
 RT zeta.";
 RL J. Immunol. 150:4364-4375(1993).
 CC -1- FUNCTION: RECEPTOR FOR THE FC REGION OF COMPLEXED IMMUNOGLOBULINS
 CC GAMMA. LOW AFFINITY RECEPTOR.
 CC -1- SUBUNIT: MAY FORM MULTISUBUNIT COMPLEX WITH OTHER HETEROPROTEINS.
 CC THIS ASSOCIATION IS REQUIRED FOR EFFICIENT CELL-SURFACE
 CC EXPRESSION. DOES NOT ASSOCIATE WITH CD3 ZETA.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST EIGHT ISOFORMS. ISOFORM A TO
 CC ISOFORM H; HAVE BEEN FOUND. THE SEQUENCE SHOWN IS THAT OF ISOFORM
 CC C. THEY MAY BE PRODUCED BY ALTERNATIVE SPLICING, BUT THE EXISTENCE
 CC OF A FAMILY OF HIGHLY HOMOLOGOUS AND PRESUMABLY DUPLICATED GENES
 CC IS ALSO POSSIBLE. THE EXPRESSION OF THE DIFFERENT ISOFORMS IS
 CC DIFFERENTIALLY REGULATED.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON NATURAL KILLERS CELLS AND
 CC MACROPHAGES.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M64368; AAA42049.1;
 DR EMBL: M64369; AAA42048.1;
 DR EMBL: M64370; AAA42050.1;
 DR EMBL: M32062; AAA41148.1;
 DR EMBL: L08446; AAA41151.1;
 DR HSSP: P12319; IALT.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 3;
 KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain; Alternative splicing.
 FT SIGNAL 1 36
 FT CHAIN 37 267
 FT DOMAIN 37 221
 FT TRANSMEM 222 241
 FT DOMAIN 242 267
 FT DOMAIN 55 111
 FT DOMAIN 136 194
 FT DISULFD 62 104
 FT DISULFD 143 187
 FT VARSPIC 145 176
 FT VARIANT 36 39
 FT ANLP -> GDLL (IN ISOFORM A AND ISOFORM H).

FT VARIANT 55 55 H)
 FT VARIANT 56 56 E -> D (IN ISOFORM A AND ISOFORM H).
 FT VARIANT 115 115 D -> G (IN ISOFORM D).
 FT VARIANT 134 134 V -> I (IN ISOFORM A AND ISOFORM B).
 FT VARIANT 138 138 E -> L (IN ISOFORM H).
 FT VARIANT 184 138 T -> R (IN ISOFORM H).
 FT VARIANT 184 184 N -> D (IN ISOFORM H).
 FT VARIANT 195 195 M -> E (IN ISOFORM H).
 FT VARIANT 254 254 K -> R (IN ISOFORM H).
 SO SEQUENCE 267 AA; 30281 MW; DD4D6A7A3F70BD92 CRC64;

Query Match 54.4%; Score 37; DB 1; Length 267;
 Best Local Similarity 54.5%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VWFHLYFLYIAV 11
 |||||
 Db 217 VWFHAFPLVM 227

RESULT 10
 C1W4_MOUSE STANDARD; PRT: 398 AA.
 AC 088454:
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE POTASSIUM CHANNEL SUBFAMILY K MEMBER 4 (TWIK-RELATED ARACHIDONIC ACID-
 STIMULATED POTASSIUM CHANNEL PROTEIN) (TRAK).
 GN KCNK4 OR TRAK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain:
 RX MEDLINE=98292450; PubMed=9628867;
 RA Fink M., Lesage F., Duprat F., Heurteaux C., Reyes R., Fosset M.,
 Lazdunski M.;
 RT "A neuronal two P domain K⁺ channel stimulated by arachidonic acid and
 polyunsaturated fatty acids."
 RL EMBO J. 17:3297-3308(1998).
 RN [2]
 RP ACTIVATION.
 RX MEDLINE=99254548; PubMed=10321245;
 RA Patel A.J., Honore E., Lesage F., Fink M., Romey G., Lazdunski M.;
 RT "Inhalational anesthetics activate two-pore-domain background K⁺
 channels."
 RL Nat. Neurosci. 2:422-426(1999).
 CC -1- FUNCTION: VOLTAGE INSENSITIVE, INSTANTANEOUS, OUTWARDLY RECTIFYING
 POTASSIUM CHANNEL. OUTWARD RECTIFICATION IS REVERSED AT HIGH
 EXTERNAL K⁺ CONCENTRATIONS.
 CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND
 2/TRAAT/PRUNATED, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SPINAL CORD AND EYE. NOT
 DETECTED IN HEART, SKELETAL MUSCLE, LIVER, LUNGS, KIDNEY AND
 TESTIS.
 CC -1- MISCELLANEOUS: ACTIVATED BY ARACHIDONIC ACID AND OTHER UNSATURATED
 FATTY ACIDS. NOT AFFECTED BY VOLATILE GENERAL ANAESTHETICS SUCH AS
 CHLOROFORM, DIETHYL ETHER, HALOTHANE AND ISOFURANE.
 CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM
 CHANNELS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; AF056492; AAC40181.1;
 DR MGD; MG1:1298234; Kcnk4.
 DR InterPro; IPRO000099;
 DR InterPro; IPRO01622;
 DR InterPro; IPRO03280;
 DR Pfam; PF02034; TWIK_channel; 1.
 DR PRINTS; PR01333; 2BOKRCHANNEL.
 KW Ionic channel; Transmembrane; Ion transport; Potassium transport;
 KM Glycoprotein; Alternative splicing.
 FT DOMAIN 1 3
 FT TRANSSEM 4 24
 FT DOMAIN 89 113
 FT TRANSSEM 119 139
 FT DOMAIN 140 171
 FT TRANSSEM 172 192
 FT DOMAIN 198 222
 FT TRANSSEM 235 255
 FT DOMAIN 256 398
 FT CARBOHYD 81 81
 FT CARBOHYD 84 84
 FT VARSPIC 63 67
 FT VARSPIC 68 398
 SO SEQUENCE 398 AA; 43051 MW; 478A834B7BAEC92 CRC64;

Query Match 54.4%; Score 37; DB 1; Length 398;
 Best Local Similarity 70.0%; Pred. No. 47;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VWFHLYFLYIA 10
 |||||
 Db 236 VWFHLYFLYIA 245

RESULT 11
 CRTI_CERNIC STANDARD; PRT: 621 AA.
 AC P48537;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE PHYTOENE DEHYDROGENASE (EC 1.3.-.-) (PHYTOENE DESATURASE).
 GN PDH1.
 GN Cercospora nicotianae.
 OS Cercospora nicotianae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
 OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cercospora.
 OX NCBI_TaxID=29003;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 18366;
 RX MEDLINE=94368091; PubMed=8085820;
 RA Ehrnschaff M., Daub M.E.;
 RT "Isolation, sequence, and characterization of the Cercospora
 nicotianae phytoene dehydrogenase gene."
 RL Appl. Environ. Microbiol. 60:2766-2771(1994).
 CC -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
 INTERMEDIARY OF PHYTOFUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
 DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
 CC -1- COFACTOR: FAD (PROBABLE).
 CC -1- PATHWAY: CAROTENOID BIOSYNTHESIS.
 CC -1- SIMILARITY: TO BACTERIAL PHYTOENE DEHYDROGENASES AND TO BACTERIAL
 METHOXYNEUROSPORENE DEHYDROGENASE (CRTI).

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

CC EMBL: U03903; AAB86988.1; -
DR InterPro: IPR000171; -
DR Pfam: PF02032; Phytoene_dh; 1.
DR PROSITE: PS00982; PHYTOENE_DH; 1.
KW Carotenoid biosynthesis; Oxidoreductase; FAD; Flavoprotein; NAD;
KW Transmembrane.
FT NP_BIND 11 44 FAD (ADP PART) (POTENTIAL).
FT TRANSMEM 536 556 POTENTIAL.
SQ SEQUENCE 621 AA; 69529 MW; 61B9EA7784963CA8 CRC64;

Query Match
Best Local Similarity 54.4%; Score 37; DB 1; Length 621;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 VFVHLYFLAVG 12
Db 601 WWSVLIVLVG 611

RESULT 12
CGZF_MOUSE STANDARD: PRT; 776 AA.
ID CGZF_MOUSE 060797; 060799;
AC P51944; 060797; 060799;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE G2/MITOTIC-SPECIFIC CYCLIN F.
GN CCNF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95284479; PubMed-7767003;
RA Obermayr F.O., Sutherland H.F., Kraus B., Fritsch A.-M.;
RT "Mouse cyclin F maps to a conserved linkage group on mouse chromosome
RT 17."
RL Mamm. Genome 6:149-150(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Croci L., Stayton C.L., Bossolasco M., Bianchi E., Corradi A.M.,
RA Pardi R., Consalez G.G.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: LIKELY TO BE INVOLVED IN THE CONTROL OF THE CELL CYCLE
CC DURING S PHASE AND G2 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC EMBL: Z47666; CA87695.1; -
DR EMBL: U20612; AAA62317.1; -
DR EMBL: U20636; AAA63152.1; -
DR HSSP: P30274; IVIN.
DR MGD: MGI:102551; Ccnf.
DR InterPro: IPR000553; -
DR InterPro: IPR001810; -
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00134; cyclin; 1.
DR PROSITE: PS00292; CYCLINS; 1.
DR PROSITE: PS05081; FBOX; 1.
KW Cyclin; Cell cycle; Cell division; Mitosis; Nuclear protein;
KW Alternative splicing.

```

```

FT DOMAIN 29 76 F-BOX.
FT VARSPIC 5 5 G-> GGCWMDGVDSYGTENGCAWTRAPVTRCASHPP
FT CONFLICT 81 81 KHGV (IN LONG ISOFORM).
FT CONFLICT 92 92 L-> P (IN REF. 2).
FT CONFLICT 92 92 Q-> E (IN REF. 2).
FT CONFLICT 132 132 T-> S (IN REF. 2).
FT CONFLICT 263 263 A-> T (IN REF. 2).
FT CONFLICT 290 290 S-> T (IN REF. 2).
FT CONFLICT 301 301 G-> V (IN REF. 2).
FT CONFLICT 364 364 T-> S (IN REF. 2).
FT CONFLICT 454 454 A-> S (IN REF. 2).
FT CONFLICT 464 464 H-> R (IN REF. 2).
FT CONFLICT 527 527 O-> R (IN REF. 2).
FT CONFLICT 567 568 GE-> ERR (IN REF. 2).
FT CONFLICT 602 602 S-> G (IN REF. 2).
FT CONFLICT 623 623 E-> R (IN REF. 2).
SQ SEQUENCE 776 AA; 86320 MW; 78639F381180B99E CRC64;

Query Match
Best Local Similarity 54.4%; Score 37; DB 1; Length 776;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 VFVHLYFLAVG 12
Db 39 VLFHKLWLVG 50

RESULT 13
YNLA_YEAST STANDARD: PRT; 123 AA.
ID YNLA_YEAST
AC P53926;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE VERY HYPOTHEICAL 15.0 KDA PROTEIN IN MSL1-RPC19 INTERGENIC REGION.
GN YNL114C OR M1934.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97245296; PubMed-9090055;
RA de Antoni A., D'Angelo M., Dal Pero F., Sartorello F., Pandolfo D.,
RA Pallavicini A., Lanfranchi G., Valle G.;
RT "The DNA sequence of cosmid 14-13b from chromosome XIV of
RT Saccharomyces cerevisiae reveals an unusually high number of
RT overlapping open reading frames."
RL Yeast 13:261-266(1997).
RN [1]
RP SEQUENCE FROM N.A.
RA EMBL: Z69382; CA93393.1; -
DR EMBL: Z71390; CA95993.1; -
DR SGD: S0005056; YNL114C.
DR Hypothetical protein; Transmembrane.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
SQ SEQUENCE 123 AA; 15042 MW; 984A6AD4328820D CRC64;

Query Match
Best Local Similarity 52.9%; Score 36; DB 1; Length 123;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VFVHLYFLAVG 10

```

Db 53 VWFLEFFFA 62

```

RESULT 14
FDHC_METTF STANDARD; PRT: 274 AA.
AC 050568;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE POTENTIAL FORMATE TRANSPORTER.
GN FDHC.
OS Methanobacterium thermoformicicum.
OC Archaea: Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacter.
OC NCBI_taxid=145262;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=2-245;
RX MEDLINE=97158688; PubMed=9006048;
RA Nolling J., Reeve J.N.;
RT "Growth and substrate-dependent transcription of the formate
  dehydrogenase (fdhCAB) operon in Methanobacterium thermoformicicum 2-
  245."
RT J. Bacteriol. 179:899-908(1997).
RL J. Bacteriol. 179:899-908(1997).
CC -1- FUNCTION: POTENTIAL FORMATE TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE ENT. FAMILY OF TRANSPORTERS (TC 2.44).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U52681; AAC44819.1;
DR InterPro; IPR000292;
DR Pfam; PF01226; Form_Nit_trans; 1.
DR PROSITE; PS01005; FORMATE_NITRITE_TP_1; 1.
DR PROSITE; PS01006; FORMATE_NITRITE_TP_2; 1.
KM Transpore; Transmembrane.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 62 82 POTENTIAL.
FT TRANSMEM 118 138 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 248 268 POTENTIAL.
SO SEQUENCE 274 AA; 29070 MW; BDD04734/BF7B1A4 CRC64;

```

Query Match 52.9%; Score 36; DB 1; Length 274;
 Best Local Similarity 33.3%; Pred. No. 50;
 Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VWFHVLFLYAVG 12
 :||: :||:|
 Db 207 WVFIFAFVAIG 218

```

RESULT 15
ID GAL7_YEAST STANDARD; PRT: 365 AA.
AC P08431; P04398;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GALACTOSE-1-PHOSPHATE URIDYLTRANSFERASE (EC 2.7.7.10).
GN GAL7 OR YBR018C OR YBR0226.
OS Saccharomyces cerevisiae (Baker's yeast).

```

```

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_taxid=4932;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89131252; PubMed=2851900;
RA Tajima M., Nogi Y., Fukasawa T.;
RT "Primary structure of the Saccharomyces cerevisiae GAL7 gene."
RT Yeast 1:67-77(1985).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95282516; PubMed=7762304;
RA Schaaf-Gerstenschnaefer I., Schindwolf T., Lehnert W., Rose W.,
  Zimmermann F.K.;
RT "Sequence and functional analysis of a 7.2 kb fragment of
  Saccharomyces cerevisiae chromosome II including GAL7 and GAL10 and a
  new essential open reading frame."
RT Yeast 11:79-83(1995).
RL [3]
RN SEQUENCE OF 1-184 FROM N.A.
RC STRAIN-CARLSBERGENSIS;
RX MEDLINE=84185433; PubMed=6715281;
RA Clifton B.A., Donelson J.E.;
RT "Sequence of the Saccharomyces GAL region and its transcription in
  vivo."
RT J. Bacteriol. 158:269-278(1984).
RL [4]
RN SEQUENCE OF 1-20 FROM N.A. AND SEQUENCE OF 1-7.
RX MEDLINE=84169499; PubMed=6324089;
RA Nogi Y., Fukasawa T.;
RT "Nucleotide sequence of the transcriptional initiation region of the
  yeast GAL7 gene."
RT Nucleic Acids Res. 11:855-856(1983).
RL -1- CATALYTIC ACTIVITY: UTP + ALPHA-D-GALACTOSE 1-PHOSPHATE =
  UDP-GALACTOSE + PYROPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE GALACTOSE-1-PHOSPHATE
  URIDYLTRANSFERASE FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X81324; CAA57105.1;
DR EMBL; M12348; AAA34627.1;
DR EMBL; X00215; CAA25039.1;
DR EMBL; Z35887; CAA84960.1;
DR EMBL; K01752; AAA34628.1;
DR PIR; S45873; XNBYUC.
DR HSSP; P09146; IHXO.
DR SGD; S0000222; GAL7.
DR InterPro; IPR000880;
DR InterPro; IPR001937;
DR Pfam; PF01087; Galp_UDP_transf; 1.
DR PROSITE; PS00117; GAL_P_UDP_TRANSF_I; 1.
KM Transferease; Nucleotidyltransferase; Galactose metabolism.
KW INT_MET 0
FT ACT_SITE 179 179 PROBABLE.
FT ACT_SITE 181 181 PROBABLE.
FT CONFLICT 10 10 H -> Y (IN REF. 3).
FT CONFLICT 57 57 P -> H (IN REF. 3).
FT CONFLICT 84 85 RL -> S (IN REF. 1).
FT CONFLICT 266 266 V -> A (IN REF. 1).
FT CONFLICT 344 344 T -> I (IN REF. 1).
SO SEQUENCE 365 AA; 42254 MW; EBA67AB72D19F3B CRC64;

```

Query Match 52.9%; Score 36; DB 1; Length 365;

Tue Jun 5 07:09:57 2001

us-09-284-107-29.rsp

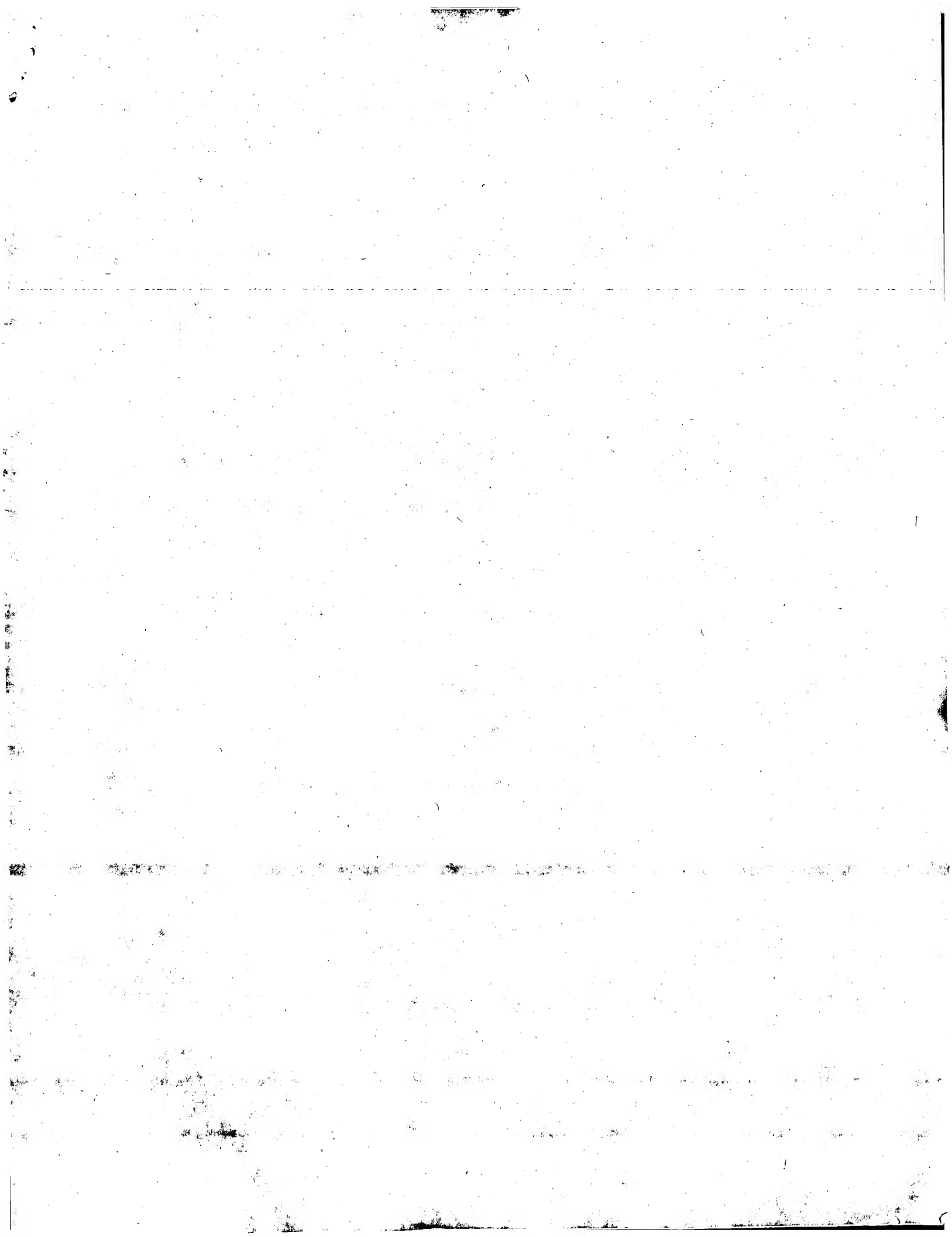
Page 9

Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 WEHVLFY 8

Db 311 WFMHMFY 317

Search completed: June 4, 2001, 12:24:09
Job time: 571 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2001, 12:23:31 ; Search time 107.68 Seconds
(without alignments)
13.062 Million cell updates/sec

Title: US-09-284-107-29
Perfect score: 68
Sequence: 1 VWFHLYFLAVG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15:
2: sp_archaea:
3: sp_bacteria:
4: sp_fungi:
5: sp_human:
6: sp_invertebrate:
7: sp_mammal:
8: sp_mhc:
9: sp_organelle:
10: sp_plant:
11: sp_rodent:
12: sp_unclassified:
13: sp_vertebrate:
14: sp_virus:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	188	4	092638
2	68	100.0	280	4	092637
3	68	100.0	374	4	092663
4	68	100.0	375	4	092495
5	63	92.6	349	6	09M270
6	45	66.2	460	2	09XD16
7	41	60.3	153	2	09WXU4
8	41	60.3	278	14	P89139
9	41	60.3	340	5	018731
10	41	60.3	1852	3	013281
11	41	60.3	1867	3	09P8W4
12	40	58.8	88	14	089405
13	40	58.8	266	2	052498
14	40	58.8	448	2	031428
15	39	57.4	168	2	09R017
16	39	57.4	378	10	09M3W5
17	39	57.4	1075	10	09LNP5
18	38.5	56.6	272	4	095004
19	38.5	56.6	1084	4	09UPP1

20	38	55.9	84	1	026862	026862 methanobact
21	38	55.9	111	10	09XEF7	09XEF7 arabidopsis
22	38	55.9	172	6	09RTU2	09RTU2 oryctolagus
23	38	55.9	197	2	09R0U2	09R0U2 streptomyces
24	38	55.9	223	2	049704	049704 mycobacteri
25	38	55.9	311	2	033142	033142 mycobacteri
26	38	55.9	314	1	058219	058219 pyrococcus
27	38	55.9	343	10	09SVI6	09SVI6 arabidopsis
28	38	55.9	348	5	09W280	09W280 drosophila
29	38	55.9	354	3	008417	008417 saccharomyc
30	38	55.9	357	10	P93565	P93565 solanum tub
31	38	55.9	391	10	09ZU52	09ZU52 arabidopsis
32	38	55.9	393	10	09SUJ4	09SUJ4 arabidopsis
33	38	55.9	395	10	09SXX4	09SXX4 nicotiana p
34	38	55.9	398	10	09SXX5	09SXX5 nicotiana p
35	38	55.9	497	5	09W1D4	09W1D4 drosophila
36	38	55.9	907	5	09VZW0	09VZW0 drosophila
37	38	55.9	1869	3	093809	093809 magnaporthe
38	37	54.4	148	8	09TH31	09TH31 campeloma p
39	37	54.4	148	8	09TH30	09TH30 campeloma p
40	37	54.4	148	8	09TH29	09TH29 campeloma p
41	37	54.4	148	8	09T414	09T414 campeloma p
42	37	54.4	148	8	09T3V8	09T3V8 campeloma p
43	37	54.4	148	8	09T3U4	09T3U4 campeloma p
44	37	54.4	190	5	09U5Q4	09U5Q4 entodinium
45	37	54.4	194	8	047579	047579 onchocerca

ALIGNMENTS

RESULT	ID	Query Match	Score	DB	Length	Pred. No.	Matches	Indels	Gaps
1	092638	100.0%	68	4	188	0.00025	12	0	0
2	092637	100.0%	68	4	280	0.00025	12	0	0
3	092663	100.0%	68	4	374	0.00025	12	0	0
4	092495	100.0%	68	4	375	0.00025	12	0	0
5	09M270	92.6%	63	6	349	0.00025	12	0	0
6	09XD16	66.2%	45	2	460	0.00025	12	0	0
7	09WXU4	60.3%	41	2	153	0.00025	12	0	0
8	P89139	60.3%	41	14	278	0.00025	12	0	0
9	018731	60.3%	41	5	340	0.00025	12	0	0
10	013281	60.3%	41	3	1852	0.00025	12	0	0
11	09P8W4	60.3%	41	3	1867	0.00025	12	0	0
12	089405	58.8%	40	14	88	0.00025	12	0	0
13	052498	58.8%	40	2	266	0.00025	12	0	0
14	031428	58.8%	40	2	448	0.00025	12	0	0
15	09R017	57.4%	39	2	168	0.00025	12	0	0
16	09M3W5	57.4%	39	10	378	0.00025	12	0	0
17	09LNP5	57.4%	39	10	1075	0.00025	12	0	0
18	095004	56.6%	38.5	4	272	0.00025	12	0	0
19	09UPP1	56.6%	38.5	4	1084	0.00025	12	0	0

AC 092637;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR I.
 GN BL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=93053454; PubMed=1430234;
 RA Porjes A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
 RA Kimberly R.P.;
 RT "Novel Fc gamma receptor I family gene products in human mononuclear
 cells";
 RT J. Clin. Invest. 90:2102-2109(1992).
 RL EMBL; L03419; AAA35825.1; -;
 DR HSSP; P12319; IALT.
 DR INTERPRO: IPR003006; -;
 DR PFAM: PF00047; 19; 2.
 DR PRODOM: PD002534; -; 1.
 SO SEQUENCE 280 AA; 32232 MW; C6C6C45AE3D345C6 CRC64;

Query Match 100.0%; Score 68; DB 4; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.00036;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHLYFLAVG 12
 Db 195 VWFHLYFLAVG 206

RESULT 3
 ID 092663 PRELIMINARY; PRT; 374 AA.
 AC 092663;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR I.
 GN AL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=93055454; PubMed=1430234;
 RA Porjes A.J., Redecha P.B., Doebele R., Pan L.C., Salmon J.E.,
 RA Kimberly R.P.;
 RT "Novel Fc gamma receptor I family gene products in human mononuclear
 cells";
 RT J. Clin. Invest. 90:2102-2109(1992).
 RL EMBL; L03418; AAA36049.1; -;
 DR HSSP; P12319; IALT.
 DR INTERPRO: IPR003006; -;
 DR PFAM: PF00047; 19; 3.
 DR PRODOM: PD002534; -; 1.
 SO SEQUENCE 374 AA; 42632 MW; D33D59398CEEA699 CRC64;

Query Match 100.0%; Score 68; DB 4; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.00048;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHLYFLAVG 12
 Db 289 VWFHLYFLAVG 300

RESULT 4
 ID 092495 PRELIMINARY; PRT; 375 AA.
 AC 092495;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR TYPE I (FC GAMMA RIB-FC GAMMA RECEPTOR).
 GN CD64, OR FC<GAMMA>RIB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Benech P.D., Sastry K.N., Iyer R.R., Eichbaum Q.G., Ravesh D.,
 RA Ezekowitz A.B.;
 RT Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-224 FROM N.A.
 RX MEDLINE=93018827; PubMed=1402657;
 RA Benech P.D., Sastry K., Iyer R.R., Eichbaum Q.G., Ravesh D.P.,
 RA Ezekowitz R.A.;
 RT "Definition of interferon gamma-response elements in a novel human Fc
 gamma receptor gene (Fc gamma Rib) and characterization of the gene
 structure";
 RT J. Exp. Med. 176:1115-1123(1992).
 RL EMBL; M91555; AAA58414.1; JOINED.
 DR EMBL; M91550; AAA58414.1; JOINED.
 DR EMBL; M91551; AAA58414.1; JOINED.
 DR EMBL; M91552; AAA58414.1; JOINED.
 DR EMBL; M91553; AAA58414.1; JOINED.
 DR EMBL; M91554; AAA58414.1; JOINED.
 DR EMBL; M91555; AAA58414.1; JOINED.
 DR EMBL; S45709; AAD13842.1; JOINED.
 DR EMBL; S45707; AAD13842.1; JOINED.
 DR EMBL; S45708; AAD13842.1; JOINED.
 DR EMBL; S45704; AAD13842.1; JOINED.
 DR EMBL; S45705; AAD13842.1; JOINED.
 DR HSSP; P12319; IALT.
 DR INTERPRO: IPR003006; -;
 DR PFAM: PF00047; 19; 3.
 DR PRODOM: PD002534; -; 1.
 SO SEQUENCE 375 AA; 42881 MW; A84D464C70DD0P91 CRC64;

Query Match 100.0%; Score 68; DB 4; Length 375;
 Best Local Similarity 100.0%; Pred. No. 0.00048;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VWFHLYFLAVG 12
 Db 290 VWFHLYFLAVG 301

RESULT 5
 ID 09M270 PRELIMINARY; PRT; 349 AA.
 AC 09M270;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE FC GAMMA RECEPTOR I.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidea;
 CC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yan Y., Li X., Wang A., Zhang G.;

RT "Molecular cloning and identification of full-length cDNA encoding
 RT high affinity Fc receptor for bovine IgG (Fc gamma RI).";
 RL Vet. Immunol. Immunopathol. 75:151-159(2000).
 DR EMBL; AF162866; AAF80477.1; -;
 KW Receptor.
 SO SEQUENCE 349 AA; 39608 MW; DOB77B2EF9408C02 CRC64;

Query Match 92.6%; Score 63; DB 6; Length 349;
 Best Local Similarity 91.7%; Pred. No. 0.0033;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VWFHLYFLYAVG 12
 |||||
 DB 289 VWFHLYFLYAVG 300

RESULT 6
 OXND16 PRELIMINARY; PRT; 460 AA;
 AC OXND16;
 DT 01-NOV-1999 (TRENBLREL. 12, Created)
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
 DE PREPROTEIN TRANSLOCASE SECY SUBUNIT.
 GN SECY.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SEROVAR LAI;
 RX MEDLINE=94014964; PubMed=8409911;
 RA Gravekamp C., van de Kemp H., Franzen M., Carrington D., Schoone G.J.,
 RA Van Eys G.J., Everard C.O., Hartskeerl R.A., Terpstra W.J.;
 RT "Detection of seven species of pathogenic leptospires by PCR using two
 RT sets of primers."
 RL J. Gen. Microbiol. 139:0-0(0).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SEROVAR LAI;
 RA Zuerner R.L., Hartskeerl R.A., van de Kemp H., Bal A.E.;
 RT "Sequence analysis of a large ribosomal protein gene operon from
 RT Leptospira interrogans."
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SECY/SEC1-ALPHA FAMILY.
 DR EMBL; AF15283; AADA0603.1; -;
 DR INTERPRO; IPR001525; -;
 DR INTERPRO; IPR002048; -;
 DR INTERPRO; IPR002208; -;
 DR PFM; PF00344; secy; 1.
 DR PRINTS; PR00303; SECYTRMLCASE.
 DR PROSITE; PS00018; EF_HAND_1; UNKNOWN_1.
 DR PROSITE; PS00095; CS_MTASE_2; UNKNOWN_1.
 DR PROSITE; PS00755; SECY_1; 1.
 DR PROSITE; PS00756; SECY_2; 1.
 DR Protein transport; Transmembrane; Translocation.
 SO SEQUENCE 460 AA; 50851 MW; C453EBB82PB566E7 CRC64;

Query Match 66.2%; Score 45; DB 2; Length 460;
 Best Local Similarity 45.5%; Pred. No. 5.3;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 VWFHLYFLYAVG 11
 :|:|:|:|:
 DB 319 VWFHLYFLYAVG 329

RESULT 7
 OXND16

ID O9WXU4 PRELIMINARY; PRT; 153 AA.
 AC O9WXU4;
 DT 01-NOV-1999 (TRENBLREL. 12, Created)
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
 DE HYPOTHETICAL 17.1 KDA PROTEIN.
 GN TM0089.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MSB / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RL Nature 399:323-329(1999).
 DR EMBL; AE001695; AAD35183.1; -;
 DR TIGR; TM0089;
 KW Hypothetical protein.
 SO SEQUENCE 153 AA; 17143 MW; 3059E8A5E96C12A1 CRC64;

Query Match 60.3%; Score 41; DB 2; Length 153;
 Best Local Similarity 66.7%; Pred. No. 9.1;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 VWFHLYFLYAVG 12
 |||||
 DB 15 VWFHLYFLYAVG 26

RESULT 8
 P89139 PRELIMINARY; PRT; 278 AA.
 ID P89139;
 DT 01-MAY-1997 (TRENBLREL. 03, Created)
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)
 DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)
 DE LARGE GLYCOPROTEIN.
 OS Simian hemorrhagic fever virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage.
 OX NCBI_TaxID=38143;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LVR 42-0/M6941;
 RA Zeng L., Godeny E.K., Methven S.L., Brinton M.A.;
 RL Virology 207:0-0(0).
 [2]
 RP SEQUENCE OF 251-278 FROM N.A.
 RC STRAIN-LVR 42-0/M6941;
 RX MEDLINE=95191057; PubMed=7884922;
 RA Godeny E.K., Zeng L., Smith S.L., Brinton M.A.;
 RT "Molecular characterization of the 3' terminus of the simian
 RT hemorrhagic fever virus genome."
 RL J. Virol. 69:2679-2683(1995).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LVR 42-0/M6941;
 RX MEDLINE=97361760; PubMed=9218721;
 RA Smith S.L., Wang X., Godeny E.K.;
 RT "Sequence of the 3' end of the simian hemorrhagic fever virus
 RT genome."
 RL Gene 191:205-210(1997).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LVR 42-0/M6941;

Q9P8W4 PRELIMINARY; PRT; 1867 AA.
 AC Q9P8W4
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE CHITIN SYNTHASE.
 GN CHS2.
 OS Blumeria graminis.
 OC Eukaryota; Fungi; Ascomycota; Erysiphales; Erysiphaceae; Blumeria.
 ON NCBI_TaxID=34373;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CC1;
 RA Zhang Z.G., Hall A., Perfect E., Gurr S.J.;
 RT "Differential expression of two chitin synthase genes of Blumeria graminis."
 RL Mol. Plant Pathol. 1:125-138(2000).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CC1;
 RA Zhang Z.G., Gurr S.J.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF189366; AAF04279.2;
 SQ SEQUENCE 1867 AA; 207129 MW; 0F36C2D219B1D9E0 CRC64;

Query Match 60.3%; Score 41; DB 3; Length 1867;
 Best Local Similarity 70.0%; Pred. No. 1e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 FHVLYFYLAVG 12
 DB 196 FHVLYFYLAVG 205

RESULT 12
 ID 089405 PRELIMINARY; PRT; 88 AA.
 AC 089405
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
 DE HYPOHETICAL 10.3 KDA PROTEIN.
 GN A70L.
 OS Paramoecium bursaria chlorella virus 1 (PCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
 ON NCBI_TaxID=10506;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9513167; PubMed=7831789;
 RA Lu Z., Li Y., Zhang Y., Kutish G.F., Rock D.L., Van Etten J.L.;
 RT "Analysis of 45 kb of DNA located at the left end of the chlorella virus PCV-1 genome."
 RL Virology 206:339-352(1995).
 DR EMBL: U42580; AAC96438.1;
 KM Hypothetical protein.
 SQ SEQUENCE 88 AA; 10265 MW; 10382DA045764EED CRC64;

Query Match 58.8%; Score 40; DB 14; Length 88;
 Best Local Similarity 66.7%; Pred. NO. 7.9;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VWFHLYFYL 9
 DB 21 IMFHLYFYL 29

RESULT 13
 ID 052498 PRELIMINARY; PRT; 266 AA.
 AC 052498

DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE O-METHYL TRANSFERASE.
 GN MERA.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bradyrhizobium group; Bradyrhizobium.
 ON NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110SPC4;
 RA Becker B.O., Kosch K., Parniske M., Muller P.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF039306; AAC04824.1;
 DR INTERPRO: IPR000051;
 DR INTERPRO: IPR001601;
 KW Transferrase.
 SQ SEQUENCE 266 AA; 29164 MW; 0318A45D2A1B958C CRC64;

Query Match 58.8%; Score 40; DB 2; Length 266;
 Best Local Similarity 85.7%; Pred. NO. 23;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 WFWHLYF 8
 DB 165 WFWHLYF 171

RESULT 14
 ID 031428 PRELIMINARY; PRT; 448 AA.
 AC 031428
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE YHDB PROTEIN.
 GN YHDB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 ON NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolydin A., Borchert S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritsch C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haelegh J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,
 RA Kodayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Laber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue V.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate I.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Solido B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsita P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wanduit R., Wedler E., Wedler H., Weltzenegger T.,

RA Winters P., Mijat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168:
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168:
 RA Haga K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.,
 RT "Sequence analysis of the 70kb region between 17 and 23 degree of the
 RT Bacillus subtilis chromosome".
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z99105; CAB11990.1;
 DR EMBL: AB006424; BAA33093.1;
 DR INTERPRO: IPR000179;
 DR PROSITE: PS00193; CYTOCHROME B_OO; UNKNOWN_1.
 SO SEQUENCE 448 AA; 51880 MW; 3870F3A47E445548 CRC64;

Query Match 58.8%; Score: 40; DB 2; Length 448;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 WPFVLFYLAIV 11
 ||:|:|:|:
 Db 296 WYILFFELIV 305

RESULT 15
 O9R017 PRELIMINARY; PRT: 168 AA.
 AC O9R017:
 DT 01-MAY-2000 (TREMREL. 13, Created)
 DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
 DT 01-JUN-2000 (TREMREL. 14, Last annotation update)
 DE HYPOTHETICAL 18.7 KDA PROTEIN (FRAGMENT).
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID-1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EGD:
 RA Fsihl H., Cossart P.;
 RT "STM applied to Listeria monocytogenes".
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF102170; AAF04741.1;
 DR INTERPRO: IPR000566;
 DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
 KW Hypothetical protein.
 FT NON_TER 1
 SO SEQUENCE 168 AA; 18718 MW; 9CF3891496C0EB8D CRC64;

Query Match 57.4%; Score 39; DB 2; Length 168;
 Best Local Similarity 41.7%; Pred. No. 22;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 WPFVLFYLAIV 12
 ||:|:|:|:
 Db 86 IWFPVMAFVAIG 97

Search completed: June 4, 2001, 12:23:32
 Job time: 595 sec